

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2003, 19:20:24 ; Search time 213.709 Seconds
(without alignments)
8277.893 Million cell updates/sec

Title: US-09-284-180A-1
Perfect score: 4008
Sequence: 1 gccagggccgcgcagtagc.....aaaaaaaaaaaaaaaaaaaa 4008

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	218	5.4	2433	4	US-09-300-958A-24
2	215.2	5.4	4157	4	US-08-556-422A-1
3	94.2	2.4	3524	4	US-09-077-940A-3
4	79.2	2.0	3692	4	US-09-077-940A-1
5	72.6	1.8	3288	2	US-08-620-694A-1
6	72.6	1.8	3288	3	US-09-022-255-1
7	72.6	1.8	3288	3	US-09-022-696-1
8	72.6	1.8	3288	3	US-08-978-773-1
9	72.6	1.8	3288	3	US-09-022-253-1
10	72.6	1.8	3288	3	US-09-022-260-1
11	72.6	1.8	3288	3	US-09-022-259-1
12	72.6	1.8	3288	3	US-09-022-257-1
13	67.4	1.7	3358	3	US-08-957-063-4
14	67.4	1.7	3358	4	US-09-487-685-4
15	67.4	1.7	3358	4	US-08-802-805D-4
16	66.6	1.7	2854	1	US-08-121-713D-57
17	66.6	1.7	2854	1	US-08-835-268-57
18	66.6	1.7	2854	2	US-09-060-692-57
19	66.6	1.7	2854	3	US-08-833-391-57
20	66.6	1.7	2854	4	US-09-060-610-57
21	66.6	1.7	2854	5	PCT-US94-10151A-57
22	66.4	1.7	3757	2	US-09-016-366A-13
23	66.4	1.7	3757	2	US-08-978-404B-19
24	66.4	1.7	80246	3	US-09-078-294-4
25	65.4	1.6	35828	4	US-09-449-218D-17
26	65.4	1.6	35828	4	US-09-668-529A-17
27	65.4	1.6	35828	4	US-09-668-037A-17

c	28	64.8	1.6	80595	3	US-09-078-294-3	Sequence 3, Appli
	29	64	1.6	3381	3	US-09-009-119-1	Sequence 1, Appli
	30	64	1.6	3381	4	US-09-371-507-1	Sequence 1, Appli
	31	64	1.6	3383	5	PCT-US95-09098-1	Sequence 1, Appli
c	32	63.8	1.6	4157	2	US-08-871-2668-1	Sequence 1, Appli
c	33	63.8	1.6	4157	2	US-08-819-458A-1	Sequence 1, Appli
c	34	63.8	1.6	4157	2	US-09-018-864A-1	Sequence 1, Appli
c	35	63.8	1.6	4157	3	US-08-871-267B-1	Sequence 1, Appli
c	36	63.8	1.6	4157	3	US-09-618-419-1	Sequence 1, Appli
	37	63.6	1.6	4771	2	US-08-866-650-2	Sequence 2, Appli
	38	63.6	1.6	4771	2	US-09-021-287-2	Sequence 2, Appli
	39	63.6	1.6	4771	3	US-09-240-473-2	Sequence 2, Appli
c	40	63.2	1.6	55298	4	US-09-491-356C-1	Sequence 1, Appli
	41	63	1.6	48974	3	US-08-920-422-17	Sequence 17, App
c	42	62.2	1.6	91	1	US-08-222-177A-107	Sequence 107, App
	43	62.2	1.6	200	1	US-08-222-177A-8	Sequence 8, Appli
c	44	61.8	1.5	3588	1	US-08-197-792-32	Sequence 32, Appli
c	45	61.8	1.5	3588	1	US-08-459-850-32	Sequence 32, Appli

ALIGNMENTS

RESULT 1

US-09-300-958A-24
; Sequence 24, Application US/09300958A
; Patent No. 6495319
; GENERAL INFORMATION:
; APPLICANT: McCrelland, Michael
; APPLICANT: Welsh, John
; TITLE OF INVENTION: Reduced Complexity Nucleic Acid Targets and Methods of
; TITLE OF INVENTION: Using Same
; FILE REFERENCE: P-PH 3457
; CURRENT APPLICATION NUMBER: US/09/300,958A
; CURRENT FILING DATE: 1999-04-27
; PRIOR FILING DATE: 60/083,331
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/098,070
; PRIOR FILING DATE: 1998-08-27
; PRIOR APPLICATION NUMBER: 60/118,624
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 2433
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-300-958A-24

Query Match	5.4%	Score 218;	DB 4;	Length 2433;
Best Local Similarity	52.1%	Pred. No. 3.4e-46;		
Matches	768;	Conservative	0;	Mismatches 640;
Indels	56;	Gaps	10;	
Qy	259	TTTGACGGCTCTCATAGCTACAATTACTCTCTCTCTGTGGATTCCTGCCCTCCACACA	318	
Db	22	TTTGAAGCTGAAACATCTCCAACATACAGGCCCTCTCTGTGTGAGCAGGATGGAAGACG	81	
Qy	319	CTTTACGTGGTGCAGGGATAGCATCTTCGCTTAA-----CCCTCCCTCTCTCT	369	
Db	82	CTGTATGTGGGGCCCGAGAGGCCCTCTTTGACATTAACAGCAACCTCAGCTTCTTGCCA	141	
Qy	370	GGGAAAGACCCCAAGGATCGACTGGATGATGCTGAGACTACACACACAACTGCAGG	429	
Db	142	GGCGGGAGTACCAGAGCTACTGTGGATGCAATGCTGCACAGACGACGAGCTGCAGC	201	
Qy	430	AAGAAAGCAAGAA---AGAGGACGAATGTACAAATTTATCCAGATTTCTGCCATGTC	486	
Db	202	TTCAAGGCAAGGACCCCAAGCGTACTGTGCTGACAAATACATCAAGATCTCTGCCACTC	261	
Qy	487	AATGCCCTCTACCTCTCAGCTGGGGACCTTCCTCTTTGATCCGAAGTGGGGGTATT	546	
Db	262	AACAGACCCACCTGCTCACCTGTGGCAGCGCCCTTCAGCCCTCTGTGTGTACATT	321	

Query Match	1.8%	Score 72.6;	DB 3;	Length 3288;
Best Local Similarity	66.0%	Pred. No. 1.2e-08;		
Matches 105: Conservative	0:	Mismatches 54:	Indels 0:	Gaps 0:

[illegible]

RESULT 7
US-09-022-696-1
; Sequence 1, Application US/09022696
; Patent No. 6072037
; GENERAL INFORMATION:
; APPLICANT: Yao, Zhengbin
; APPLICANT: Spriggs, Melanie
; APPLICANT: Fanslow, William
; TITLE OF INVENTION: No. 6072037el Receptor That Binds IL-17
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2003, 17:51:04 ; Search time 983.19 Seconds
(without alignments)
11004.333 Million cell updates/sec

Title: US-09-284-180A-1
Perfect score: 4008
Sequence: 1 gcgcagccgcgcagtagc.....aaaaaaaaaaaaaaaaaaaa 4008

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03:*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*

2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*

3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*

4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*

5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*

6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*

7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*

8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*

9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*

10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*

11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*

12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*

13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*

14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*

15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*

16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*

17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*

18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*

19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*

20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*

21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*

22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*

23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*

24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4008	100.0	4008	19	AAV07279 Rat semaphorin W e
2	3122.2	77.9	4024	21	AA287977 Mouse semaphorin W
3	2331	58.2	2331	19	AAV07280 Rat semaphorin W e
4	1875.2	46.8	2971	22	AAF93776 Human cDNA encodin
5	1673.6	41.8	2893	20	AA886126 DNA encoding SBSEM
6	1673.6	41.8	2894	20	AA886127 EST sequence for D
7	1404.2	35.0	2315	19	AAV07281 Human semaphorin W
8	1391.2	34.7	1761	19	AAV07282 Human semaphorin W

9	654	16.3	799	20	AA86128	EST sequence for D
10	597.8	14.9	928	25	ABT32087	NOVX DNA sequence
11	527.4	13.2	777	25	ABT32088	NOVX DNA sequence
12	505.2	12.6	671	24	ABL89791	Human polynucleoti
13	408	10.2	669	22	AAF93972	Primer specific fo
14	251.8	6.3	3293	22	AAD08048	Human extracellular
15	248.6	6.2	2155	22	AAC84887	Human SEC6 nucleic
16	248.6	6.2	2156	22	AAC84892	Human SEC6 nucleic
17	248.6	6.2	2284	22	AAC84888	Human SEC7 nucleic
18	248.6	6.2	3556	22	AAD08283	Human secreted pro
19	248.6	6.2	3776	24	ABN83983	Human gene sequenc
20	247.8	6.2	2558	24	AAD28948	Human MOL5b cDNA.
21	247.8	6.2	3112	24	AAD28949	Human MOL5c cDNA.
22	247	6.2	3781	21	AA337092	Human PRO1480 (UNQ
23	247	6.2	3781	22	AA346151	Human DNA encoding
24	247	6.2	3781	22	AAF54381	Primer #75 used in
25	247	6.2	3781	25	ACA57909	Human PRO1480 cDNA
26	247	6.2	3781	25	ABX98379	Human cDNA encodin
27	247	6.2	3781	25	ABX98881	Novel human secret
28	247	6.2	3781	25	ACA05926	Human secreted/tra
29	247	6.2	3781	25	ABX97970	Human PRO polynucl
30	247	6.2	3781	25	ABX78754	Human PRO polynucl
31	247	6.2	3781	25	ABX75767	Human cDNA encodin
32	247	6.2	3781	25	ABX76972	Human PRO polynucl
33	247	6.2	3781	25	ABX16812	Human cDNA encodin
34	246	6.1	3868	24	AAD28947	Human MOL5a cDNA.
35	245.4	6.1	3503	22	AAF29461	Murine M-Sema-F CD
36	245.4	6.1	3766	24	ABV77918	Hypoxia-induced pr
37	245.4	6.1	3766	24	ABN83984	Human gene sequenc
38	245.4	6.1	3766	25	ACC51059	Human bladder canc
39	245.4	6.1	3766	25	ABX76369	Lung cancer-associ
40	238.4	5.9	8095	24	ABX92031	Lung specific nucl
41	238.4	5.9	8144	25	AA113397	Breast specific re
42	238.2	5.9	2703	21	AAA47437	Sequence encoding
43	238	5.9	333	19	AAV07286	Human semaphorin W
44	228	5.7	2390	24	ABQ11109	DNA encoding human
45	224.4	5.6	2615	24	ABQ99272	Human coding sequ

ALIGNMENTS

RESULT 1

AAV07279

ID AAV07279 standard; cDNA to mRNA; 4008 BP.

XX AAV07279;

AC AAV07279;

XX 08-SEP-1998 (first entry)

DT Rat semaphorin W encoding cDNA with 5'UTR and 3'UTR.

DE Rat semaphorin W; nerve extension inhibitor; antiallergic; anticancer;

XX Immunosuppressant; gene therapy; diagnosis; research reagent; ds.

XX Rattus norvegicus.

OS Rattus norvegicus.

XX Key Location/Qualifiers

PH 5'UTR 1..75

FT CDS /*tag= a

FT CDS /*tag= b

FT CDS /*tag= "semaphorin W"

FT 5'UTR 2407..3977

FT polyA_signal 3978..4008

FT /*tag= d

XX WO9815628-A1.

XX 16-APR-1998.

XX 03-OCT-1997; 97WO-JP03549.

XX 09-OCT-1996; 96JP-0287636.
XX (SUMU) SUMITOMO PHARM CO LTD.
XX Kikuchi K, Kimura T;
XX WPI; 1998-261015/23.
XX P-PSDB; AAW51313.
XX Nerve extension inhibitor protein semaphorin W - is useful as
XX therapeutic drug and diagnostic and research reagent
XX Claim 2; Page 55-57; 90pp; Japanese.
XX The present sequence encodes rat semaphorin W. Semaphorin W and its
XX derivatives are nerve extension inhibitors which are useful as
XX anti-allergic, immunosuppressant and anticancer agents. The DNA
XX encoding semaphorin W can also be used in gene therapy, e.g. using
XX a viral vector. The proteins, peptides, DNA and antibodies which
XX recognise the protein or peptides, can be used as diagnostic or
XX research reagents. Semaphorin W can be used as a screen for
XX semaphorin W antagonists with possible therapeutic use.
XX Sequence 4008 BP; 775 A; 1135 C; 1141 G; 957 T; 0 other;
XX
XX Query Match 100.0%; Score 4008; DB 19; Length 4008;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 4008; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCGAGGCCGCGCAGTAGCGGTACTAAGTAGAGGCTGCTGGACGCGCCGCCCGCGCGCT 60
DB 1 GCCGAGGCCGCGCAGTAGCGGTACTAAGTAGAGGCTGCTGGACGCGCGCCGCCCGCGC 60
QY 61 CAGGGGGAGCCAGAGAGCTTCCAGAGGCCGAGCGCGCGCGCGCGCGCGCGCGCT 120
DB 61 CAGGGGGAGCCAGAGAGCTTCCAGAGGCCGAGCGCGCGCGCGCGCGCGCGCGCT 120
QY 121 CCGGCTCTTCCCTTCCCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
DB 121 CCGGCTCTTCCCTTCCCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 181 CCGGCTGCGGGCGCGCTCCCGCGCTCAGTGCCAGAACCTGCTGCTGCTGCTGCTGCTGCT 240
DB 181 CCGGCTGCGGGCGCGCTCCCGCGCTCAGTGCCAGAACCTGCTGCTGCTGCTGCTGCTGCT 240
QY 241 GACTCCTATCTACACCGGTTTCCAGCGTCTCATAGTACAAATTAATCTGCTGCTGCTGCTG 300
DB 241 GACTCCTATCTACACCGGTTTCCAGCGTCTCATAGTACAAATTAATCTGCTGCTGCTGCTG 300
QY 301 GATCCTGCTCCACACACTTTTACGTCGCTGCGGAGTAGCATCTTCCGTTTAAACCTC 360
DB 301 GATCCTGCTCCACACACTTTTACGTCGCTGCGGAGTAGCATCTTCCGTTTAAACCTC 360
QY 361 CCGTTCTCTGGGAAAGACCCGAGGATCGACTGGATGGTACCTGAGACTCACAGACAG 420
DB 361 CCGTTCTCTGGGAAAGACCCGAGGATCGACTGGATGGTACCTGAGACTCACAGACAG 420
QY 421 AACTCAGGAAGAAAGGCAAGAAAGAGGACGAAATGTCACAAATTTATCCAGATTCGCG 480
DB 421 AACTCAGGAAGAAAGGCAAGAAAGAGGACGAAATGTCACAAATTTATCCAGATTCGCG 480
QY 481 ATTGTCAATGCTCTACCTCTCAGTCCGCGACCTTCGCTTTTATCCGAGTGC 540
DB 481 ATTGTCAATGCTCTACCTCTCAGTCCGCGACCTTCGCTTTTATCCGAGTGC 540
QY 541 GTTATTGATGTCCAGTTTCCAGAGGTTGAAGACTTGAGAGCGCGCGGGGAAATGT 600
DB 541 GTTATTGATGTCCAGTTTCCAGAGGTTGAAGACTTGAGAGCGCGCGGGGAAATGT 600
QY 601 CTTTTTGAGCCAGCTCAACGGTTCAGACGCTGTAATGGCTGGGGGCTCTCTACACGCC 660
DB 601 CTTTTTGAGCCAGCTCAACGGTTCAGACGCTGTAATGGCTGGGGGCTCTCTCTACACGCC 660

QY 661 ACTGTGAAGAACTTCTCTGGGACTGAGCCATCATCTCCGAGCTGTGGGTGAGCTGAG 720
DB 661 ACTGTGAAGAACTTCTCTGGGACTGAGCCATCATCTCCGAGCTGTGGGTGAGCTGAG 720
QY 721 GACTGGATTCGAACAGAGACCTTGTCTATCTCGCTTAATGTCCAGCTTTGTCCAGCT 780
DB 721 GACTGGATTCGAACAGAGACCTTGTCTATCTCGCTTAATGTCTCCAGCTTTGTCCAGCT 780
QY 781 ATGGTCTCTGAGCCAGCTGAGTGGGGGATGAAGATGGAGATGAATCTTTTTC 840
DB 781 ATGGTCTCTGAGCCAGCTGAGTGGGGGATGAAGATGGAGATGAATCTTTTTC 840
QY 841 TTCACGAGACCTCCGAGCTGTGACTCTATGAGCGCATCAAGTCCCAAGAGTGGCC 900
DB 841 TTCACGAGACCTCCGAGCTGTGACTCTATGAGCGCATCAAGTCCCAAGAGTGGCC 900
QY 901 CGAGTGTGCGGGGAGCTTGGGGCAGAGACCTTTCAGCAGATGAGACGCTTT 960
DB 901 CGAGTGTGCGGGGAGCTTGGGGCAGAGACCTTTCAGCAGATGAGACGCTTT 960
QY 961 CTGAAGGCTGACCTGCTGCTGCGCCAGGCGCGGAGCATGGCCGGCTCTGCGAG 1020
DB 961 CTGAAGGCTGACCTGCTGCTGCGCCAGGCGCGGAGCATGGCCGGCTCTGCGAG 1020
QY 1021 GCTATGGCAGAGCTTTCGGCTCAGCTTGAGCGGGAACCCCATCTTTTATGGGATCTTT 1080
DB 1021 GCTATGGCAGAGCTTTCGGCTCAGCTTGAGCGGGAACCCCATCTTTTATGGGATCTTT 1080
QY 1081 TCCTCCCAGTGGGAAGAGCTGCCATCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
DB 1081 TCCTCCCAGTGGGAAGAGCTGCCATCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
QY 1141 CCGGCGAGTGTGAATGGTCCCTTTAGAGAGCTTAAACATGACTGCAACAGGGAGCTGCT 1200
DB 1141 CCGGCGAGTGTGAATGGTCCCTTTAGAGAGCTTAAACATGACTGCAACAGGGAGCTGCT 1200
QY 1201 GTCATGGACAACAGAGTGTCCCGAGCCAGACCTGGAGAGTGTGATGCAACAACTGAAG 1260
DB 1201 GTCATGGACAACAGAGTGTCCCGAGCCAGACCTGGAGAGTGTGATGCAACAACTGAAG 1260
QY 1261 CTCAGCAGTGTGGATTCCTACTCTCCCTGCGAGACCGGGTGTCTCACTTTATCAGAGAC 1320
DB 1261 CTCAGCAGTGTGGATTCCTACTCTCCCTGCGAGACCGGGTGTCTCACTTTATCAGAGAC 1320
QY 1321 CACCTCTCATGAGCAGCGCGGTGTCCGGCTGAGCGCGCGCGCTGCTGCTGCTACTACA 1380
DB 1321 CACCTCTCATGAGCAGCGCGGTGTCCGGCTGAGCGCGCGCGCTGCTGCTGCTACTACA 1380
QY 1381 GATACAGCTTATCTCAGAGTCTGTGCGCCACAGGGTGTGACAGCTCTCTCAGGGAAAGATAT 1440
DB 1381 GATACAGCTTATCTCAGAGTCTGTGCGCCACAGGGTGTGACAGCTCTCTCAGGGAAAGATAT 1440
QY 1441 GACGTGTCTACCTGGGACAGAGATGACACTTCCACCGGGCTGTGGGCTTTGGAGCT 1500
DB 1441 GACGTGTCTACCTGGGACAGAGATGACACTTCCACCGGGCTGTGGGCTTTGGAGCT 1500
QY 1501 CAGCTCAGTGTCTGGAGATCTGGCTTGTTCAGAGACACACCGGTTGAGAGCATG 1560
DB 1501 CAGCTCAGTGTCTGGAGATCTGGCTTGTTCAGAGACACACCGGTTGAGAGCATG 1560
QY 1561 AAATTTGTACACAGATTTGCTCTGCTGGCTCCCATCTGAGGTGACAACTGACACCC 1620
DB 1561 AAATTTGTACACAGATTTGCTCTGCTGGCTCCCATCTGAGGTGACAACTGACACCC 1620
QY 1621 AGCAACTGTGGCGCTCTCCAGAGTGTCTCGGAGTGTATCTTGGCCAGAGACCCCGTGTGC 1680
DB 1621 AGCAACTGTGGCGCTCTCCAGAGTGTCTCGGAGTGTATCTTGGCCAGAGACCCCGTGTGC 1680
QY 1681 GCCTGGAGCTTCCGGCTTTCATGCTGTGTGCGCCAGCGGAGACACCGGGGATGGTT 1740
DB 1681 GCCTGGAGCTTCCGGCTTTCATGCTGTGTGTGCGCCAGCGGAGACACCGGGGATGGTT 1740

QY 1741 CAAGATATAGAGTCAGCGGATGTCTCTCTCTTTGTGTCCAAAAGAACCTGGAGACATCCC 1800
Db 1741 CAAGATATAGAGTCAGCGGATGTCTCTCTTTGTGTCCAAAAGAACCTGGAGACATCCC 1800
QY 1801 GTAGTGTTTGAAGTTCGGGTGGCTACTGTGGGCCACGTGGCTGCGCATGTTCGCCAGT 1860
Db 1801 GTAGTGTTTGAAGTTCGGGTGGCTACTGTGGGCCACGTGGCTGCGCATGTTCGCCAGT 1860
QY 1861 TCTGCTGGGCATCCTGTGTGGCACCAGCCAGTGGAGTGAAGTGGCTCACATCCCGG 1920
Db 1861 TCTGCTGGGCATCCTGTGTGGCACCAGCCAGTGGAGTGAAGTGGCTCACATCCCGG 1920
QY 1921 AGGATGACATAGAGTGGTGGTGACCCAGGGGCCATGGGGCTTATGTCTGGAGTGT 1980
Db 1921 AGGATGACATAGAGTGGTGGTGACCCAGGGGCCATGGGGCTTATGTCTGGAGTGT 1980
QY 1981 CAGAGGGTGGAGCGCCCGCTGGTGGCTGCTTATAGCTTGGTGGGGCAGCCAGGG 2040
Db 1981 CAGAGGGTGGAGCGCCCGCTGGTGGCTGCTTATAGCTTGGTGGGGCAGCCAGGG 2040
QY 2041 GGACCCCTAAACCGGGCCACACCGTTGTGGGGCTGGATTGGTGGCTTCTCTGGGT 2100
Db 2041 GGACCCCTAAACCGGGCCACACCGTTGTGGGGCTGGATTGGTGGCTTCTCTGGGT 2100
QY 2101 GTTCTGAGCATCCCTCACTCTCCTCTGATTGGTCGCGTCAGCAGCGTCGGCAGAG 2160
Db 2101 GTTCTGAGCATCCCTCACTCTCCTCTGATTGGTCGCGTCAGCAGCGTCGGCAGAG 2160
QY 2161 AGGAGCTTCTAGCTAGAGACAAGTGGGCTTAGATCTGGGGCTCCACCTTCTGGGACC 2220
Db 2161 AGGAGCTTCTAGCTAGAGACAAGTGGGCTTAGATCTGGGGCTCCACCTTCTGGGACC 2220
QY 2221 ACAAGCTATAGTCAGGACCTCCCTCTCTCGCTGAAGATGAAGGCTGCCCTGSGC 2280
Db 2221 ACAAGCTATAGTCAGGACCTCCCTCTCTCGCTGAAGATGAAGGCTGCCCTGSGC 2280
QY 2281 CTGGGTAAAGCGGGCAGTGGTTTGGTGGCTTCCCTCCACCTTCTCGTGGATTCGTG 2340
Db 2281 CTGGGTAAAGCGGGCAGTGGTTTGGTGGCTTCCCTCCACCTTCTCGTGGATTCGTG 2340
QY 2341 CCAAGCCAGCCACATCCGGCTCAGTGGGGCGCTCTAGCCAGCTGTGATGACACCTCC 2400
Db 2341 CCAAGCCAGCCACATCCGGCTCAGTGGGGCGCTCTAGCCAGCTGTGATGACACCTCC 2400
QY 2401 ATCTAAAGCCGGGAAATGACTGCCACCATGAGCAGTCTCTGGAATAGTGGCTACCA 2460
Db 2401 ATCTAAAGCCGGGAAATGACTGCCACCATGAGCAGTCTCTGGAATAGTGGCTACCA 2460
QY 2461 AGACCATGATCATGGCTGCTCCTTCTCTTGGAGTCTGTGTTCACACATTAAGTGTG 2520
Db 2461 AGACCATGATCATGGCTGCTCCTTCTCTTGGAGTCTGTGTTCACACATTAAGTGTG 2520
QY 2521 TCCTCTGGACCTGGACCTGGCTTGGCCAGATTCCCTGATTCTCATGAGATCAACCT 2580
Db 2521 TCCTCTGGACCTGGACCTGGCTTGGCCAGATTCCCTGATTCTCATGAGATCAACCT 2580
QY 2581 GTAACCTTCTCGCATGGCTCTTGTCTTGGGCCCATCAGCTTGTGGGTGGAGTAAGGAC 2640
Db 2581 GTAACCTTCTCGCATGGCTCTTGTCTTGGGCCCATCAGCTTGTGGGTGGAGTAAGGAC 2640
QY 2641 ATAGCCCCGGAAAGGGAATCAGTGTGGAGTGTGGGGCTGTGTGCCCTGGCTCCTT 2700
Db 2641 ATAGCCCCGGAAAGGGAATCAGTGTGGAGTGTGGGGCTGTGTGCCCTGGCTCCTT 2700
QY 2701 GTGGTGGCTGTATGATTTCCAGTCTGCTGACTCTGGGGAGCGCATGATCCCTGACTGC 2760
Db 2701 GTGGTGGCTGTATGATTTCCAGTCTGCTGACTCTGGGGAGCGCATGATCCCTGACTGC 2760
QY 2761 CTTGAGATCTCTCCCACTCAGTTTCCCTTGGCTTGGGAAGAGTGTGTCTATACACTG 2820
Db 2761 CTTGAGATCTCTCCCACTCAGTTTCCCTTGGCTTGGGAAGAGTGTGTCTATACACTG 2820
QY 2821 GTGTGCCCTAGAAGGCCCTGTCCATGTGTGATGAGCAGACAGGGCGGCTCGCTCGCTTT 2880

Db 2821 GTGTGCCCTAGAAGGCCCTGTCCATGTGTGATGAGACAGGGCGGCTGCTCGGTGCTTT 2880
QY 2881 TGGGAGTTCGCGAGAGAAAGTTCGAATGGGGGCAACTTAACCTCGGTAGCCAGTGG 2940
Db 2881 TGGGAGTTCGCGAGAGAAAGTTCGAATGGGGGCAACTTAACCTCGGTAGCCAGTGG 2940
QY 2941 GAAACCCACATGCCCGTCCCATCACCCACAGCGCTCTTTAACTTTTTCAGCAAAAGTTCCC 3000
Db 2941 GAAACCCACATGCCCGTCCCATCACCCACAGCGCTCTTTAACTTTTTCAGCAAAAGTTCCC 3000
QY 3001 AAAGTACCTTCTGGTGGGAAGGCAAGAGATGTGGCCCGCTTCTTCTCTCTCTCTCT 3060
Db 3001 AAAGTACCTTCTGGTGGGAAGGCAAGAGATGTGGCCCGCTTCTTCTCTCTCTCTCTCT 3060
QY 3061 TTTTCCCTTCTGGCTGCCAACCACTGCCGTGCCACCGCTTCCCTGGCTGGAGTGG 3120
Db 3061 TTTTCCCTTCTGGCTGCCAACCACTGCCGTGCCACCGCTTCCCTGGCTGGAGTGG 3120
QY 3121 AGGCTGAGTCTCTCTCTCTCTCTCTTCCATTTTAAATGAACCTTCAACAATTTTAAATAT 3180
Db 3121 AGGCTGAGTCTCTCTCTCTCTCTCTTCCATTTTAAATGAACCTTCAACAATTTTAAATAT 3180
QY 3181 GGGGATGACAAATGACTTTTTTCCCCAGAAAAGTGTAGGAAATACAAGCAGGTTAAA 3240
Db 3181 GGGGATGACAAATGACTTTTTTCCCCAGAAAAGTGTAGGAAATACAAGCAGGTTAAA 3240
QY 3241 GAAATTTGCTCCTCAGTGACTTTTCCCTTAAAGCAGAGTCCCTCAGCTAGCGTC 3300
Db 3241 GAAATTTGCTCCTCAGTGACTTTTCCCTTAAAGCAGAGTCCCTCAGCTAGCGTC 3300
QY 3301 TGTGACTCCCTGAAATGTATGCTCTGTGGACTTCCCTGAAATGTATGCAAAAGTGTCT 3360
Db 3301 TGTGACTCCCTGAAATGTATGCTCTGTGGACTTCCCTGAAATGTATGCAAAAGTGTCT 3360
QY 3361 GT 3420
Db 3361 GT 3420
QY 3421 GT 3480
Db 3421 GT 3480
QY 3481 TATAGTACCACACTTGGGCCACATGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3540
Db 3481 TATAGTACCACACTTGGGCCACATGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3540
QY 3541 GGCTGCTGAGGCCAGCCCTGGACATTTGCTGTGGAGAGCCCAATCTCTCCAGAA 3600
Db 3541 GGCTGCTGAGGCCAGCCCTGGACATTTGCTGTGGAGAGCCCAATCTCTCCAGAA 3600
QY 3601 CTTGGAAGCTAGGCTTGGCGGTGTGAAAGCAGCCACATCTCCCTTCTTGTGTCTATTC 3660
Db 3601 CTTGGAAGCTAGGCTTGGCGGTGTGAAAGCAGCCACATCTCCCTTCTTGTGTCTATTC 3660
QY 3661 TCCTGT 3720
Db 3661 TCCTGT 3720
QY 3721 CCTGGTGGGGACCTCAAAACCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3780
Db 3721 CCTGGTGGGGACCTCAAAACCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3780
QY 3781 CTGCTTCTCTTGACAGCAGCTGTGAATCTACTCAGAGTCCCTTGGTGTGGAGTTCCTCG 3840
Db 3781 CTGCTTCTCTTGACAGCAGCTGTGAATCTACTCAGAGTCCCTTGGTGTGGAGTTCCTCG 3840
QY 3841 GTGGCTTTGAGTAGGATCTTTTGGGTGGCATCTAACCTTAGCAGCATTCATCGTTCATGT 3900
Db 3841 GTGGCTTTGAGTAGGATCTTTTGGGTGGCATCTAACCTTAGCAGCATTCATCGTTCATGT 3900
QY 3901 AAAGTGGGGATATACCTACCTCAGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3960

Db	1264	AAGTCCAGCAGTTTGGATCCTCACTCTCCCTGCGACAGCGTGTGCTCACCTTTATCAGA	1323	2398	TCCATCTAAAGCCGGGAAATGACTGCCAGCCATGAGCAGTCTCTGGAAGTCTGCTA	2457	
Qy	1318	GACCACCTCTCTATGGACAGCCCGTGTTCOCGGCTCAGCGCGCCCTGCTGGTCACT	1377	Db	2404	TCATCTTAAGCC - AGGAAATGACTGCCAGCCATGAGCAATCTCTGGAAGTCTGCTA	2462
Db	1324	GACCACCGCTCATGGACAGCCCGTGTTCAGCTGATGGCGCCCTGCTGGTCACT	1383	Qy	2458	CCAGAGCC-----ATGATCATGGCTGCTCCTTTCTCTTTGGAGTCT	2497
Qy	1378	ACAGATACAGCCTATCTCAGAGTGTGGCCGCACAGGGTGACCAAGCCTCTCAGGGAAGAA	1437	Db	2463	CGAAGACAGGGTCACCTGAACCTGGGAGAGACATCATGGCTGCTCCATCTCTTTGGAGTCT	2522
Db	1384	ACGATACAGCCTATCTCAGATTTGGCTCACCGGTCACCGGTCACAGCCCTCTCAGGGAAGAA	1443	Qy	2498	GTGTGTTTACACATTTAGTGTCTGCTCTCTGACCTGAGCTGAGCTGAGGGCCAGACCTTTG	2547
Qy	1438	TATGACGTGCTCTACCTGGGACAGAGATGGACACCTCTCACCGGGCTGTGCGCATGGA	1497	Db	2523	GTGTGATCACACGCTCAGTGTCTGTCTCTGGACCTGATGAGCTGAGGGCCAGACCTTTG	2582
Db	1444	TATGATGTCTCTACCTGGGACAGAGATGGGACCTCTCATCGGGCTGTGCGCATCGGA	1503	Qy	2548	CCAGATTCCTGATTTCTCATGAGAGATCAACCTCTTAACCTTCTCGGATGGCTCTTTGTCT	2607
Qy	1498	GCTCAGCTCAGTGTCTGGAGGATCTGGCCTGTGTCCAGAACACACAGCCGTTGACAGC	1557	Db	2583	CCTGATTCCTGATTTCTCATGAGAGATCAACCTGTAGCCCTCTCTGATGGCTCTTTGTCT	2642
Db	1504	GCTCAGCTCAGTGTCTGGAGGATCTGGCCTGTGTCCGGAAACACAGACCGGTTGAAGC	1563	Qy	2608	TGGGCC---CATCAGCTTTGGGGTGGAGTAAGACATAGG-----	2645
Qy	1558	ATGAAATTTACCAACGATTTGGCTCTGGTGGGCTCCCATACTAGAGGTGACACAGTGAAC	1617	Db	2643	TGGCCCATCATCAGCTTTGGGGTGGAGCAAGACACAGGCTCTGATTTTGTCTTTATGGTG	2702
Db	1564	ATGAAATTTGACCATGATTTGGCTTCTGGTGGGCTCCCATACTAGAGGTGACACAGTGAAC	1623	Qy	2646	--CCCGCGAAGGGAATCAGTGTGGAGGTAGTTGGGGCGTGTGCGCTGCGCTCCCTGT	2702
Qy	1618	ACCAGCAACTGTGGCCCTCTCCAGAGCTGCTCGAGTGTATCTGGGCCAGGACCCCGTG	1677	Db	2703	GGCCCGCGAAGGAGGAGCAGCTCGGGGTGGTTGGAGCGGTGTGCCCTGCATCCCGTGT	2762
Db	1624	ACCAGCAACTGTGGCCCTCTCCAGAGCTGCTCAGAAATGTATCTGGGCCAAGGATCCCGTG	1683	Qy	2703	GGTGGCTGT---ATGATTTTCCAGTCTGCTGACTCTCTGGGAGCGCATGATCCCCGTGACTG	2759
Qy	1678	TGCGCCTGGAGCTTCCCGCTTGATGCTTGTGTGGCCACAGCGCGAGACCGCGGATG	1737	Db	2763	GGTGGCTGTATGATGATTTTCCAGTCTGCTGACTCTCTGGGAGCGCATGATCCCCGACTG	2822
Db	1684	TGTGCTTGGAGCTTCCCGCTCGATGCTTGTGTGGCCCATGAGCGGAGCACCCTGGGATG	1743	Qy	2760	CCTTGAGATCTCTCCCAACTCAGTTTCCCTTG---CTCTGGAAGAGTGTGTCTATACAC	2818
Qy	1738	GTTCAAGATATAGAGTACAGCGGATGCTCTCTTCTTGTGTCCAAAGAACCTTGGAGAAT	1797	Db	2823	CCTTGAGATCTCTCCCAACTCAGTTTCCCTTGTCCTTTGAAGAGTGTGTCTCTATACAC	2882
Db	1744	GTTCAAGACATAGAGTACAGAGATGCTCTCTTCTTGTGTCCAAAGAACCTTGGAGAAT	1803	Qy	2819	TGCTGCTCCTAGAGGCTGTCCATGTGTGATGACGACAGACGCGCTGCTCGGTGCT	2878
Qy	1798	CCGTAGTGTTTGAATTCGGTGGCTACTGTGGGCCAGGTGGTCTTGCCATGTTCCCC	1857	Db	2883	TGGTGTGCTCGAAGACCTGTCCAGGTGTGTGGATGACAGGCGCGGTGCTAGGTGCT	2942
Db	1804	CCGTAGTGTTTGAATTCGGTGGCTACGCTGGGTGGGCCACGTTGTCCTGCTCCCTCC	1863	Qy	2879	TTTGGGAGTCTGGAGAGAAAGTTTGGATGGGAGCAACTTTAACCCCTCGGTAGCACATGA	2938
Qy	1858	AGTCTGCTGGGACCTCTGTGTGGCACAGCCAGTGGAGTGGCTCAGTCTCCCTCC	1917	Db	2943	TTTGGGAGTCTGGAGAGAAAGTTTGGATGGGAGCAACTTTAACCCCTCGGTAGCACATGA	3002
Db	1864	AGTCTGCTGGGACCTCTGTGTGGCACAGCCAGTGGAGTGGCTCAGTCTCCCTCC	1923	Qy	2939	GGGAACACATGCCCCGTGCTCCCATCACCCACAGCGCTCTTTTAACTTTGAGCAAGTTC	2998
Qy	1918	CGGAGGATGACTTAGAGTGGTGGTACCCCCAGGGGCCATAGGGGGCTTATGCTTGGAG	1977	Db	3003	GGGAACACCA---CGACCCCATCACCCACAGCGCTCTTTCAACTTTGAGCAAGTTC	3058
Db	1924	CGAAGGATGGCTTAGAGTGGTGGTACCCCCAGGGGCCATAGGGGGCTTATGCTTGGAG	1983	Qy	2999	CCAAAGTACCTTCTGGTGGGAGGAGGAGCAGCAGACATGTGGCCCGCTCTCTCCTTG	3058
Qy	1978	TGTACAGAGGTGGAGCGCCCGCTGTGTGGCTCTTATAGCTTGGTGTGGGAGCCAG	2037	Db	3059	CCAAAGTACCTTCT---GTTGGGAGGAGCAGCAGCAGCTGG---CCCGTCTCTCTCTCA	3116
Db	1984	TGTACAGAGGTGGAGCGCCCGCTGTGTGGCTCTTATAGCTTGGTGTGGGAGCCAG	2043	Qy	3059	TCCTTCTCTCTGCTGCCAACCACTGCGCTGGCCACCGCTGCGCTTTTCCCTGGTGGAGT	3118
Qy	2038	CGGGGACCTCAAAACCGGGCCACACCGTTGTGGGGCTGGATTTGGCTTTCTCTG	2097	Db	3117	TCCTTCTCTCTGCTGCC---ACCACCTGCGGTGCCACCGCTCTTTTCCCTGGTGGAGT	3175
Db	2044	AGGGGGCCGCAAAACCGGGCCACACCGTTGTGGGGCTGGCTGGCTTTTCTCTG	2103	Qy	3119	GGAGGCTGAGTCTCTGCTCTGTTTCCATTTTAAATGAACTTTCACACATTTCTAATA	3178
Qy	2098	GGTGTCTTGAGCATCCCTCACTCTCTCTCTGATTTGGTGGCCGTGAGCAGCTGGCGA	2157	Db	3176	GGAGGCTGAGCCCTCTTTCTCTGCTTCCATTTTAAATGAACTTTCACACATTTCTAGTA	3235
Db	2104	GGTGTCTTGAGCATCTCTCACTCTCTCTCTGATTTGGTGGCCGTGAGCAGCGCGCGA	2163	Qy	3179	TTGGGGATGACAAATGACTTTTTCCTCCGAGAAAGTGTGTAGGAAA---TACAAGCAGT	3237
Qy	2158	CAGAGGAGCTTCTAGCTAGACACAAGTGGGCTTAGATCTGGGGCTCCACCTTCTGGG	2217	Db	3236	TTGAGGATGACAAAAG-----GAACCCGAGAAAGTGTGTAGGAAATTTACACAGGTT	3289
Db	2164	CAGAGGAGCTTCTAGCTAGACACAAGTGGGCTTAGATCTGGGGCTCCACCTTCTGGG	2223	Qy	3238	AAAGAAGATTTGCTCAGTGAATTTTCCCTTTGCCCCCTAAAGCAGGAGTCCCTCAGCTAGC	3297
Qy	2218	ACCACAGCTATAGCTAGGACCCCTCCCTCTCTCTCTGCTGAGATGAACGCTGCCCTG	2277	Db	3290	CAAGAAGATTTGCTCAGTGAATTTTCCCTTTGCCCCCTAAAGCAGGAGTCCCTTAGC---	3345
Db	2224	ACCACAGCTATAGCTAGGACCCCTCCCTCTCTCTCTGCTGAGATGAACGCTGCCCTG	2283	Qy	3298	GTCTGTGAGTCTCCTGAAATTTGATGCTGTGGACTCTCTGAAATTTGATGCAAGTGTG	3357
Qy	2278	GCCCTGGGTAAAGGGGACAGTGGTTTGGTGGCTTCCCTCCACCTTCTCTGCTGATCT	2337	Db	3346	-----AGGACTCTCCTGAAATTTGATG---CAAGAAATTTGCTGTCTCCTATGTGGCATGTG	3397
Db	2284	GCCCTGGGTAAAGGGGACAGTGGTTTGGTGGCTTCCCTCCACCTTCTCTGCTGATCT	2343	Qy	3358	TCGT	3417
Qy	2338	TGCCCAAGCCAGCCCAATCCCGCTCAGTGGGGCGCTCTAGCCAGCTGTGTGTGTGTGTGT	2397	Db	3398	TGCGTGGCTGT	3442
Db	2344	TGCCCGAGCCAGCCCAATCCCGCTCAGTGGGGCTCTCTGTGGCCAGTGTGATGAGACG	2403				

Db 721 GCTGAGTGGGGGATGAAGATGGAGACGATGAATCTTTTCTTACGGAGACCTCC 780
QY 856 CGAGTGTGGACTCCTATGAGCGCATCAAGGTCCTCCAGAGTGGCCGAGTGTGTGGGGG 915
Db 781 CGAGTGTGGACTCCTATGAGCGCATCAAGGTCCTCCAGAGTGGCCGAGTGTGTGGGG 840
QY 916 GACCTTGGGGGACAGAACCCCTTACGAGAGATGGACGACGTTTCTGAAGGTGACCTG 975
Db 841 GACCTTGGGGGACAGAACCCCTTACGAGAGATGGACGACGTTTCTGAAGGTGACCTG 900
QY 976 CTGTGCCAGGGCCGAGCATGCGGGGCTCCGGGTTCTGAGGCTATGGCAGACTT 1035
Db 901 CTGTGCCAGGGCCGAGCATGCGGGGCTCCGGGTTCTGAGGCTATGGCAGACTT 960
QY 1036 CGGCTCAGCTGAGCGGGAACCCCATCTTTTATGGGATCTTTTCTCCAGTGGAA 1095
Db 961 CGGCTCAGCTGAGCGGGAACCCCATCTTTTATGGGATCTTTTCTCCAGTGGAA 1020
QY 1096 GGAGTGCCTATCTGTGTGTGTGCTTCCGACCCCAAGACATPCGGGGCAGTGTCTGAAT 1155
Db 1021 GGAGTGCCTATCTGTGTGTGTGCTTCCGACCCCAAGACATPCGGGGCAGTGTCTGAAT 1080
QY 1156 GGTCCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCTGTCTATGGACAAGAG 1215
Db 1081 GGTCCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCTGTCTATGGACAAGAG 1140
QY 1216 GTGCCCCAGCCAGACTGGAGAGTGCATCGCAACAAATGAAGCTCCAGCAGTTTGGAA 1275
Db 1141 GTGCCCCAGCCAGACTGGAGAGTGCATCGCAACAAATGAAGCTCCAGCAGTTTGGAA 1200
QY 1276 TCCTCACTCTCCTGTGCGACACCGGTGCTCACCTTTATCAGAGACCACTCTCATGGAC 1335
Db 1201 TCCTCACTCTCCTGTGCGACACCGGTGCTCACCTTTATCAGAGACCACTCTCATGGAC 1260
QY 1336 AGGCGCTGTTCGGGCTGACGGCGCCCTGCTGTCTACTACAGATACAGCCTATCTC 1395
Db 1261 AGGCGCTGTTCGGGCTGACGGCGCCCTGCTGTCTACTACAGATACAGCCTATCTC 1320
QY 1396 AGAGTCTGTGCCCCACAGGGTGACAGGCTCTCAGGAGAAAGATATGACGTCTACCTG 1455
Db 1321 AGAGTCTGTGCCCCACAGGGTGACAGGCTCTCAGGAGAAAGATATGACGTCTACCTG 1380
QY 1456 GGGACAGAGATGACACCTTCCACCGGGCTGTGCGCATTTGGAGCTCAGTCTAGTGTG 1515
Db 1381 GGGACAGAGATGACACCTTCCACCGGGCTGTGCGCATTTGGAGCTCAGTCTAGTGTG 1440
QY 1516 GAGGATCTGGCTTGTTCACAGACACACCGGTTGAGAGCATGAATTTGTACCAGAT 1575
Db 1441 GAGGATCTGGCTTGTTCACAGACACACCGGTTGAGAGCATGAATTTGTACCAGAT 1500
QY 1576 TGGCTCTGTGGCTCCCATCTAGGTGACACAACTGAACACCACTGTGGCGGT 1635
Db 1501 TGGCTCTGTGGCTCCCATCTAGGTGACACAACTGAACACCACTGTGGCGGT 1560
QY 1636 CTCAGAGCTGTGCGAGTGTATCTCTGGCCAGGACCCCGTGTGCGCTGGAGCTTCCGG 1695
Db 1561 CTCAGAGCTGTGCGAGTGTATCTCTGGCCAGGACCCCGTGTGCGCTGGAGCTTCCGG 1620
QY 1696 CTTGATGCTTGTGTGGCCACGCGCGGAGACCGGGGATGGTTCAAGATATAGATCA 1755
Db 1621 CTTGATGCTTGTGTGGCCACGCGCGGAGACCGGGGATGGTTCAAGATATAGATCA 1680
QY 1756 GCGGATGCTCTTCTTTGTGTCCAAAGAACCTGGAGAACATCCCGTGTGAGGTT 1815
Db 1681 GCGGATGCTCTTCTTTGTGTCCAAAGAACCTGGAGAACATCCCGTGTGAGGTT 1740
QY 1816 CCGGTGGCTACTGTGGCCACGCTGTCTCTGCTTCCCGCAGTCTTGTGCTGGCATCC 1875
Db 1741 CCGGTGGCTACTGTGGCCACGCTGTCTCTGCTTCCCGCAGTCTTGTGCTGGCATCC 1800
QY 1876 TGTGTGTGGACACAGCCAGTGTGAGTGTGCTGCTCACTCCCGGAGGATGAGCTAGAG 1935
Db 1801 TGTGTGTGGACACAGCCAGTGTGAGTGTGCTGCTCACTCCCGGAGGATGAGCTAGAG 1860

QY 1936 GTGTTGTGTACCCAGGGGCTTATGCTTGCAGTGTTCAGAGGGTGGAGCC 1995
Db 1861 GTGTTGTGTACCCAGGGGCTTATGCTTGCAGTGTTCAGAGGGTGGAGCC 1920
QY 1996 GCCCGGTGTGTGGCTTATGCTTGTGGGGCAGCAGCGGGGACCTCAAAACGG 2055
Db 1921 GCCCGGTGTGTGGCTTATGCTTGTGGGGCAGCAGCGGGGACCTCAAAACGG 1980
QY 2056 GCCCACACCTGTGGGGCTGATGTTGGCTTCTCTCTGGTGTCTTGACAGATCC 2115
Db 1981 GCCCACACCTGTGGGGCTGATGTTGGCTTCTCTCTGGTGTCTTGACAGATCC 2040
QY 2116 CTCACTCTCTCTGATTTGCTCGCGTCAGCAGCGTCGGCGACAGAGGGAGCTTCTAGCT 2175
Db 2041 CTCACTCTCTCTGATTTGCTCGCGTCAGCAGCGTCGGCGACAGAGGGAGCTTCTAGCT 2100
QY 2176 AGAGACAAGTGGGCTTAGATCTGGGGCTCCACCTTCTGGGACCAAGCTATAGTCAG 2235
Db 2101 AGAGACAAGTGGGCTTAGATCTGGGGCTCCACCTTCTGGGACCAAGCTATAGTCAG 2160
QY 2236 GACCCTCCCTCTCTTCGCGCTGAAGATGAACGGCTGCCCTGGCCCTGGGTAAAGCGGGC 2295
Db 2161 GACCCTCCCTCTCTTCGCGCTGAAGATGAACGGCTGCCCTGGCCCTGGGTAAAGCGGGC 2220
QY 2296 AGTGGTTTGTGGCTTCCCTCCACCTTCTCTGCTGGATTCTTGCCTCAAGCCAGCCAC 2355
Db 2221 AGTGGTTTGTGGCTTCCCTCCACCTTCTCTGCTGGATTCTTGCCTCAAGCCAGCCAC 2280
QY 2356 ATCCGGCTCACTGGGGGCTCTAGCCACGCTGTGATGAGACCTCCATCTAA 2406
Db 2281 ATCCGGCTCACTGGGGGCTCTAGCCACGCTGTGATGAGACCTCCATCTAA 2331

RESULT 4

AAF93776
ID AAF93776 standard; cdNA; 2971 BP.

XX AC AAF93776;

XX DT 23-MAY-2001 (first entry)

XX Human cDNA encoding a membrane or secretory protein clone PSEC0074.

DE Human; secretory protein; membrane protein; vaccine; gene therapy;
KW rheumatoid arthritis; diabetes; ss.

XX Homo sapiens.

OS EPI067182-A2.

PN 10-JAN-2001.

XX 07-JUL-2000; 2000EP-0114090.

XX 08-JUL-1999; 99JP-0194179.

PR 11-JAN-2000; 2000JP-0118775.

PR 02-MAY-2000; 2000JP-0183766.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;

DR WPI; 2001-093989/11.

XX P-PSDB; AAB88349.

XX Nucleic acids encoding secretory proteins/membrane proteins, useful in
PT gene therapy or as candidate target molecules in drug development.

XX Claim 1; SEQ ID 65; 609pp + CD ROM; English.

XX This invention relates to nucleic acid sequences AAF93744 - AAF93916
CC which encode human secretory or membrane proteins represented by

Db	1753	CAAGACATAGATCAGCAGATGTCCTCTTTGTGTCCTAAAGAGCCTGGAGAACGTCCA	1812
Qy	1801	GTAGTGTTTGAAGTTCCTGGGTACTGTGTGGCCAGCTGGTCTCGCATGTTTCCCCAGT	1860
Db	1813	GTAGTGTTTGAAGTTCCTGGGTACAGCTGCGCATGTGTCTTGCATGTTCTCCAAGC	1872
Qy	1861	TCGTCTGGGCATCCTGTGTGTGGCACACAGCCAGTGGAGTGACTGGGTCACTCCCCGG	1920
Db	1873	TCAGCATGGGCATCCTGTGTGTGGCACACAGCCAGTGGAGTGACTGCACTCCGCCCG	1932
Qy	1921	AGGGATGGACTAGAGTGTGTGTGACCCACAGGGCCATGGGGCTTATGCTTGCAGTGT	1980
Db	1933	CGGGATGGACTGAGGTGTGTGTGACCCACAGGGCCATGGGGCTTATGCTGTGAATGT	1992
Qy	1981	CAGAGGGTGGAGCCGCCCGCTGGTGCCTTATAGCTTGTGTGGGCAGCCAGCGG	2040
Db	1993	CAGAGGGTGGGCAGCCCATGTGTAGCAGCTTACAGCTTGTATGGGCAGCCAGCGA	2052
Qy	2041	GGACCTCAAAACGGGGCCACACCGTGTGTGGGGCTGGATGTGTTGGCTTTCCTGGGT	2100
Db	2053	GATGCTCCGAGCCGGGCCACAC- --AGTGGGGCGGGACTGGCTGGCTTCTTTGGGG	2109
Qy	2101	GTTCCTTGCAGCATCCCTCACTCCTCCTGATGTGTGCGCGTCAGCAGCGTGGCCACAG	2160
Db	2110	ATTCTCGACATCCCTGACTCTCATCTGATTGTGCGGGCTCAGCAGCGAGGCGACAG	2169
Qy	2161	AGGGAGCTTCTAGCTAGACAAAGTGGGCTTAGATCTGGGGCTCCACCTTCTGGGACC	2220
Db	2170	AGGGAATCTGGCTAGACAAAGTGGGCTTGGACCTGGGAGCTTCCCTTGGGACC	2229
Qy	2221	ACAAGCTATAGTCAGGACCCCTCCTCTCCTTCCGCTGAAGATGAACGGTGGCCCCGGCC	2280
Db	2230	ACAAGCTACAGCCAAGACCCCTCCCTCCCTCCCTGAAAGTGAAGGTGGCGGTGGCC	2289
Qy	2281	CTGGGTAAAGGGGGCAGTGGTTTGTGTGGCTTCCCTCCACCCCTTCTGCTGGATTCTTGC	2340
Db	2290	CTGGCAAGAGGGGCAGTGGCTTTGTGTGAATCTCACACCCCTTCTGCTGTATCTTGC	2349
Qy	2341	CCAAGCCAGCCACATCCGGCTCACTGGGGCGCCCTTAGCCACGTGTGATGAGACCTCC	2400
Db	2350	CCAAGCCAGCCACATTCGGCTAACTGGGGCTCCTTAGCCACATGTGATGAACATCC	2409
Qy	2401	ATCTAAAGCGGGGAAATGACTGCCAGCCATGAGCAGTCTCTTGGAAC	2448
Db	2410	ATCTAGAGCTGGGCAATGACCACCTAGTGTATTAAGTATGATCAGCTGGAAC	2457

RESULT 5

RESULT 5
RAY96136

AAX86126
ID AAX86126 standard; DNA; 2893 BP.

0
2
4
6
8
XX
XX

AA
AC
AA86126.

AC AAX86126;
vv

XX
DT 15-SEP-1999 (first entry)

DE DNA encoding SBSEM1, a semaphorin family polypeptide.

XX	SESEN1, semaphorin; neurodegeneration; spinal injury; neuropathy;
KW	neuromuscular disorder; muscular dystrophy; psychiatric disorder;
KW	inflammatory disorder; developmental malformation;
KW	immune system disorder; cancer; viral infection; vaccine; ss.

XX

OS Homo sapiens.

XX
CO NOVOBARTENS:

XX
DN ED0324CE-21

PN EP933425-A1.

XX	1
XX	2
XX	3
XX	4
XX	5
XX	6
XX	7
XX	8
XX	9
XX	10
XX	11
XX	12
XX	13
XX	14
XX	15
XX	16
XX	17
XX	18
XX	19
XX	20
XX	21
XX	22
XX	23
XX	24
XX	25
XX	26
XX	27
XX	28
XX	29
XX	30
XX	31
XX	32
XX	33
XX	34
XX	35
XX	36
XX	37
XX	38
XX	39
XX	40
XX	41
XX	42
XX	43
XX	44
XX	45
XX	46
XX	47
XX	48
XX	49
XX	50
XX	51
XX	52
XX	53
XX	54
XX	55
XX	56
XX	57
XX	58
XX	59
XX	60
XX	61
XX	62
XX	63
XX	64
XX	65
XX	66
XX	67
XX	68
XX	69
XX	70
XX	71
XX	72
XX	73
XX	74
XX	75
XX	76
XX	77
XX	78
XX	79
XX	80
XX	81
XX	82
XX	83
XX	84
XX	85
XX	86
XX	87
XX	88
XX	89
XX	90
XX	91
XX	92
XX	93
XX	94
XX	95
XX	96
XX	97
XX	98
XX	99
XX	100

PD 04-AUG-1999.

XX

PF 30-SEP-1998;

XX
JULY 1960,
VOL 38 NO 7

30-777-1000

PR 30-JUL-1998;
DD 30-JAN-1998;
DD 30-JAN-1998;

PR 30-JAN-1998; 98EP-0300693.

28-JUL-1998; 98GB-0016423.
(SMIK) SMITHKLINE BEECHAM PLC.
Doe TR, Hayes PD, Michalovich D;
WPI; 1999-407151/35.
P-PSDB; AAY23873.
New semaphorin family polypeptide useful for treating spinal injury
and muscular dystrophy
Claim 7; Page 13-14; 29pp; English.
The present sequence encodes SBSEMN1, a semaphorin family polypeptide.
The polypeptide may be used to screen for agonists or antagonists
including antibodies. Measurement of the level of SBSEMN1 protein
and detection of a mutation in its polynucleotide may also be used to
diagnose a disease or condition or susceptibility to a disease related
to altered expression or activity of SBSEMN1. These diseases and
conditions include neurodegeneration, spinal injury, neuropathies,
neuromuscular disorders, muscular dystrophy, psychiatric disorders,
inflammatory disorders, developmental malformations, disorders of the
immune system, cancer and viral infections. The protein may also be
used to treat these diseases and conditions by administration as a
vaccine.
Sequence 2893 BP; 579 A; 825 C; 842 G; 647 T; 0 other;

28-JUL-1998; 98GB-0016423.
(SMIK) SMITHKLINE BEECHAM PLC.
Doe TR, Hayes PD, Michalovich D;
WPI; 1999-407151/35.
P-PSDB; AAY23873.
New semaphorin family polypeptide useful for treating spinal injury
and muscular dystrophy

Qy	601	CCTTTTGAGCCAGCTCAACGGTCAGCAGCTGTAAATGGCTGGGGCGTCTCTACACGCC	660
Db	521	CTTTTGAGCCAGCTCAGCGGTCAGCAGCTGTAAATGGCTGGGGGGTCTCTATGCTGCC	580
Qy	661	ACTGTGAAGACTTCCTTGGGACTGAGCCCATCATCTCCGAGCTGTGGTCTGAGCTGAG	720
Db	581	ACTGTGAAGACTTACCCTGGGACGGAGCCAAATATACAGAGCAGTGGGTCTGCGCAG	640
Qy	721	GACTGGATTCGAACACAGACCTTGCTCATCTCTGGCTTAATGCTCCAGGCTTTGTGCGCAGCT	780
Db	641	GACTGGATTCGGACAGATACCTTGCCCTTCTGGCTGAACGCCCGCAGCTTTGTGCGCAGCC	700
Qy	781	ATGTCCTGAGCCCGCAGCTGAGTGGGGGATGAAGATGGAGACGATGAAATCTTTTTTTC	840
Db	701	GTGGCTTGTAGCCCGAGCGAATGGGGGATGAAGATGGAGACGAGAAATCTACTTCTTC	760
Qy	841	TTACGGAGACCTCCCGAGTGTGGACTTCCTATGACGGCATCAAGGTCCCAAGATGGCC	900
Db	761	TTTACGGAGACTTCCCGAGCATTTGACTCATACAGCGCATTAAGTCCCACGGTGGCC	820
Qy	901	CGAGTGTGTGGGGGGACCTTGGGGGCGAGGAAGACCCTTCAGCAGAGATGGACAGCTTT	960
Db	821	CGTGTGTGGGGGGACCTCGGGGGCGGAGACCTCCAGCAGAGATGGACAGCTTT	880
Qy	961	CTAAGGCTGACCTGTGTGTGCCAGGGCCGAGAGATGCCGGGCTCCGGGTTCTGCAG	1020
Db	881	TTGAAAGCTGACCTGTCTGTCTCCAGGGCTGAGCATGGCGGGCTCCAGTGTCTGCAG	940
Qy	1021	GCTATGCGAGACTTCGGCCTCAGCTTGGAGCGGAACCCCATCTTTATGGGATCTTT	1080
Db	941	GATGTGTGTGTCTGACCTTGAGCTTGGGGCAGGGACTCCCATCTTTATGGCATCTTT	1000
Qy	1081	TCCCTCCAGTGGGAAGAGCTGCCATCTCTGTCTGTGTGCTTCCGACCCCAAGACATC	1140
Db	1001	TCCTCCAGTGGGAGGGGCTACTATCTCTGTCTGTGTGCTTCCGACCCACAGACATT	1060
Qy	1141	CGGCGAGTGTGAATGTTCCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCCT	1200
Db	1061	CGGACAGTGTGAATGTTCCCTTTCAGAGAACTAAACATGACTGCAACAGGAGCTGCCT	1120
Qy	1201	GTCATGACACAGAGTTGCCCGACCCAGACCTGGAGAGTGCATCGCCACACACATGAAG	1260
Db	1121	GTCGTGACAAATGATGTGCCCGACCCAGACCTGGAGAGTGCATCACCACACACATGAAG	1180
Qy	1261	CTCCAGAGATTGGATCCTCACTCTCCTGCGACAGCCGGTGTCTACCTTTATCAGAGAC	1320
Db	1181	CTCCGGCAGTTTGGCTCATCTCTCCTGCTGACCGGTACTCACCCTTCATCCGGGAC	1240
Qy	1321	CACCTCTCATGGAGCGCCGTTTCCGGCTGACGGCGCCCGCTGCTGGTCACTACA	1380
Db	1241	CACCCACTCATGGACAGCCAGCTTTTCCAGCTGATGGCCAGCCCTGCTGGTCACTACA	1300
Qy	1381	GATACAGCCTATCTCAGAGTCTGGGCCACAGGTTGACACCTCTCAGGGAAGAATAT	1440
Db	1301	GATACAGCCTATCTCAGAGTCTGGGCCACAGGTTGACACCTCTCAGGGAAGAATAT	1360
Qy	1441	GACCTGTCTACCTTGGGACAGAGATGGACACCTCCACCGGCTGCGCATTTGGAGCT	1500
Db	1361	GATGTGCTTACCTTGGGACAGAGATGGACACCTCCACCGAGCAGTGGCGATCGGAGCT	1420
Qy	1501	CAGCTCAGTGTCTTGGAGTCTGGCCCTTGTTCAGAAACACACAGCCGTTGAGAGCATG	1560
Db	1421	CAGCTCAGCGTTCTTGAAGATCTGGCCCTATTCCAGAGCCACAGCCAGTTGAGAACATG	1480
Qy	1561	AAATTGTACAGATTTGGTCTCTGGTGGCTTCCCATACTGAGGTGACACAAGTGAACACC	1620
Db	1481	AAATTGTACACAGCTGGCTCTGTTGGCTCCCGTACTGAGGTGACACAAGTGAATACA	1540
Qy	1621	AGCAACTGTGCGCTTCCAGAGCTGCTCGGAGTGTATCTTGGCCCGAGGACCCGCTGTGC	1680
Db	1541	ACCAACTGTGGCGCTTCCAGAGCTGCTCAGAGTGCATCTTGGCCCGAGGACCCAGTCTGT	1600
Qy	1681	GCCTGGAGCTTTCGGGCTTGATGCTTGTGTGGCCCGACCGCCGCGAGCACCGCGGATGGTT	1740

RESULT 6	
AAx86127	
ID	AAx86127 standard; DNA; 2894 BP.
XX	
XX	AAx86127;
XX	
XX	AC
XX	
DT	15-SEP-1999 (first entry)
XX	
DE	EST sequence for DNA encoding SBSEMN1.
XX	
XX	SBSEMN1; semaphorin; neurodegeneration; spinal injury; neuropathy;
KW	neuromuscular disorder; muscular dystrophy; psychiatric disorder;
KW	inflammatory disorder; developmental malformation; EST;
KW	expressed sequence tag; immune system disorder; cancer;
KW	viral infection; vaccine; ss.
XX	
OS	Homo sapiens.
XX	
PN	EP933425-Al.
XX	
XX	
PD	04-AUG-1999.

```
XX 30-SEP-1998; 98EP-0203287.
XX PF
XX 30-JUL-1998; 98GB-0016676.
XX PR
XX 30-JAN-1998; 98EP-0300693.
XX PR
XX 28-JUL-1998; 98GB-0016423.
XX PR
XX (SMIK ) SMITHKLINE BEECHAM PLC.
XX PA
XX Doe TR, Hayes PD, Michalovich D;
XX PI
XX WPT; 1999-407151/35.
XX DR
XX P-PSDB; AAY23874.
XX DR
XX New semaphorin family polypeptide useful for treating spinal injury
XX PT and muscular dystrophy
XX PS
XX Claim 18; Page 14-15; 29pp; English.
XX
XX The present sequence represents an expressed sequence tag (EST)
CC sequence, from which DNA encoding SBSEMN1, a semaphorin family
CC polypeptide, is derived. The polypeptide may be used to screen
CC for agonists or antagonists including antibodies. Measurement
CC of the level of SBSEMN1 protein and detection of a mutation in
CC its polynucleotide may also be used to diagnose a disease or
CC condition or susceptibility to a disease related to altered
CC expression or activity of SBSEMN1. These diseases and
CC conditions include neurodegeneration, spinal injury, neuropathies,
CC neuromuscular disorders, muscular dystrophy, psychiatric disorders,
CC inflammatory disorders, developmental malformations, disorders of the
CC immune system, cancer and viral infections. The protein may also be
CC used to treat these diseases and conditions by administration as a
CC vaccine.
XX
XX Sequence 2894 BP; 580 A; 824 C; 841 G; 646 T; 3 other;
SQ
Query Match 41.8%; Score 1673.6; DB 20; Length 2894;
Best Local Similarity 82.2%; Pred. No. 0;
Matches 2012; Conservative 0; Mismatches 319; Indels 117; Gaps 3;
QY 1 GCCGAGGCGCGCAGTAGCGGTACTAAGTAGAGGCTGCTGGACGCGCGCCACCCGCGAC 60
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
35 GCCGAGGCGCGTAGCCCGGGGCGCTGAGCAGAGGCGGTAGCTGGCGCGCGACCCGCGGC 94
QY 61 CAGCGGAGGCGAGAGTGTGCGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 120
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
95 CAGCGGAGGCGAAGATGCGCGCTCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 154
QY 121 CCGGTCTTTCCCTTCCCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
155 ACAGCCTCGCGCTTCCC- - - - -GCTACTGCTGCTGGCGGTGCTGAGCGGC 199
QY 181 CCGGTGCGCGCGCGCTCCCGCTCAGTCCGAGACCTCGCTGCCCATCTCCGAGCT 240
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
200 CCGGTATCCGCGCGCGCTCCCGCTCGGTGCCAGAACCTCGCTTCCAATCTCTGAGGCT 259
QY 241 GACTCTCTATCTACCCCGGTGCGAGCGTCTCATACGTACAAATTAATCTGCTCTCTG 300
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
260 GACTCTCTCTACCCCGGTGCGAGTCCCTCACACATACATTAATCTGTTCTCTCTG 319
QY 301 GATCTCGCTCCACACATTTAGTGGGTGCGAGGATAGCATCTTCGCTTTAAACCTC 360
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
320 GATCTCGCTCCACACATTTATGTTGGCGCGCGGACACCATCTTCGCTTTATCCCTG 379
QY 361 CCCTTCTCTGGGAACACCCCGAAGGATCGACTGGATGTACTGTAGACTCAGACAG 420
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
380 CCCTTCTCAGGGAGAGACCCCGAGGATGACTGGATGGTCTCTGAGGCTCAGACAG 439
QY 421 AACTGCAAGGAAGCAAGAAAGAGGACGAATGTCACAATTTATCCAGATTTCTCGCC 480
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
440 AACTGTAGGAAGAAAGCAAGAAAG- - - - - 466
QY 481 ATTGTCAATGCCTCTACCTCTCAGTGGCGGACCTTCGCTTTTGATCCGAAAGTGGCGG 540
```

```
Db 467 ----- 466
QY 541 GTTATTGATGTGTCAGTTCACAGTTCAGAGACTTGAAGACTTCAGAGCGCGCGGGAATGT 600
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
467 -----GATGTGTCAGGTTCCAGAGCTTGAAGACTTGAGAGTCGCGCGGGAATGT 520
QY 601 CCTTTTGAGCAGCTCAACGCTCAGCAGCTGTAATGCTGGGGCGTCTCTACACCGCC 660
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
521 CCTTTTGAGCAGCTCAGCGGTGAGCAGCTGTAATGCTGGGGGTCTCTATGCTGTC 580
QY 661 ACTGTGAAGAACTTCCTGGGAGCTAGACCCCATCATCTCCGAGGTGTGGGTGCGAGTGAG 720
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
581 ACTGTGAAGAACTACCTGGGAGCGGAGCAATATACACAGAGCAGTGGGTGCTGCGGAG 640
QY 721 GACTGGATTGCAACAGAGACCTTGTCTATCTGCTTAACTCCAGCTTAACTGCTCCAGCTT 780
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
641 GACTGGATTGCGACAGATACCTTTCCTTCCCTGGAACGCCCGCAGCTTTGTCGAGCC 700
QY 781 ATGCTCTGAGCCAGCTGAGTGGGGGATGAAGATGGAGACGATGAAATCTTTTTC 840
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
701 GTGCGCTTGAGCCCGCAGCGAATGGGGGATGAAGATGGAGACGACGAAATCTACTTCTTC 760
QY 841 TTCACGAGAGCTCCCGAGTGTGGACTCTCTATGAGCGCATCAAGGTCCCAAGAGTGGCC 900
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
761 TTTACGGAGACTTCCCGAGCATTTGACTCATACAGCGCATTTAAAGTCCCAAGGTGGCC 820
QY 901 CGAGTGTGTCGGGGGACCTTGGGGCAGGAAGACCTTCAGCAGAGATGAGCAGCTTT 960
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
821 CGTGTGTGTCGGGGGACCTTCGGGGCGGAAGACCTTCAGCAGAGATGAGCAGCTTT 880
QY 961 CTGAAGGCTGACCTGTGTGTCGCCAGGCGCGAGCATGGCGGGCTCCGGGGTTCGAG 1020
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
881 TTGAAAGCTGACCTGTCTGTCCAGGSCCTGAGCATGGCGGGCTCCAGTGTCTCGAG 940
QY 1021 GCTATGCGAGAGCTTCGCGCTCAGCCTGGAGCGGGAACCCCATCTTTTATGGATCTTT 1080
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
941 GATGTGCTGTGCTTTCGACCTGAGCTTGGGGCAGGGACTCCCATCTTTTATGGCATCTTT 1000
QY 1081 TCCTCCAGTGGGAAGGAGCTGCCATCTGTGTGTGTGCTTCCGACCCCAAGACATC 1140
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1001 TCCTCCAGTGGGAGGGGCTACTATCTGTGTGTGTGCTTCCGACCCCAAGACATC 1060
QY 1141 CGGCGAGTGTGATGTCTCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCCT 1200
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1061 CGGACAGTGTGATGTCTCTTCAGAGAACTAAACATGACTGCAACAGAGGACTGCCT 1120
QY 1201 GTCATGACAACAGAGTGGCGCCAGCCAGACCTGGAGAGTGCATGCGCCACAACATGAAG 1260
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1121 GTCGTGACAATGATGTGCCCGCAGCTGAGAGAGTGCATCACAACAACATGAAG 1180
QY 1261 CTCAGCAGTTTGGATCTCCTCTCCCTGCCAGACCGCTGCTCACCCTTTATCAGAGAC 1320
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1181 CTCGGGACCTTGGCTCATCTCTCTCCCTGCTGACCGGCTACTCACCCTTCATCCGGGAC 1240
QY 1321 CACCCTCTCATGGACAGGCGCGTGTCCCGGCTGACGCGCGCGCCCTGCTGGTCACTACA 1380
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1241 CACCCACTCATGGACAGGCGCAGTGTTCAGCTGATGGCCACCCCTGCTGGTCACTACA 1300
QY 1381 GATACAGCTTATCTCAGTCTGGGCGCCACAGGTTGACAGCCTCTCAGGGAAGAATAT 1440
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1301 GATACAGCTTATCTCAGATCGTGGCGCCACAGGTTGACAGCCTCTCAGGGAAGAATAT 1360
QY 1441 GACTGTCTTACCTGGGACAGAGGATGGACCTCCACCGGGCTGTGGCATTGGAGCT 1500
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1361 GATGTGCTTACCTGGGACAGAGATGGACCTCCACGAGGAGTGGGATCGGAGCT 1420
QY 1501 CAGCTCAGTGTCTTGGAGGATGTGGCCTTGTTCAGAACCCACAGCGGTTGAGAGCATG 1560
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1421 CAGCTCAGGCTTCTTGAAGATCTGGCCTTATTTCCACAGCCACAGCCAGTTGAGACATG 1480
QY 1561 AAATTGTACACAGATTTGGCTCTCGGTGGGTCCCATCTACTGAGGTGACACAGTGAACACC 1620
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



```
Db 421 CCCATCTTTATGCGACTCTTTTCTTCCAGGTGGGAGGGGCTACTATCTCTGCTGTCTGT 480
Qy 1120 GCCTTCGACCCCAAGACATCCGGGCACTGCTGAATGGTCCCTTTAGAGAGCTAAAACAT 1179
Db 481 GCCTTCGACCCCAAGACATTCGACAGTGTCTGAATGGTCCCTTCAGAGAACTAAAACAT 540
Qy 1180 GACTGCAACAGGGGAGCTGCTGTCTCATGGACAAGAGTGGCCCAAGCCAGACCTTGGAGAG 1239
Db 541 GACTGCAACAGAGAGCTGCTGTCTGTGGACAAATGATGTGCCCAAGCCAGACCTTGGAGAG 600
Qy 1240 TGCATCCGACACACATGAAGCTCCAGCAGTTTGGATCCCTCACTCTCCCTGCCAGACGC 1299
Db 601 TGCATCCGACACACATGAAGCTCCGACCTTTTGGCTCATCTCTCTCCCTGCCAGACGC 660
Qy 1300 GTGCTCACTTTATFACAGACACCACTCTCATGACAGAGCCGCTGTGTCGCGGCTGACGGC 1359
Db 661 GTACTCACTTATCCCGGACCACTCATGACAGAGCCAGTGTTCAGCTGATGGC 720
Qy 1360 CGCCCTCTGCTGTCTACAGATACAGCTATCTCAGAGTCGTGGCCCAAGAGGTGACC 1419
Db 721 CACCCCTGTGTGTCTACTACAGATACAGCTATCTCAGAGTCGTGGCCCAAGAGGTGACC 780
Qy 1420 AGCTCTCAGGGAAGATATGACGTCTCTACCTGGGACAGAGATGGACACTCCAC 1479
Db 781 AGCTCTCAGGGAAGAGTATGATGTCTCTACCTGGGACAGAGATGGACACTCCAC 840
Qy 1480 CGGCTGTGGCATTTGGAGCTCAGCTCAGTGTCTTTGGAGGATCTGGCTTTGTTCCAGAA 1539
Db 841 CGASCAGTCGGATCGGAGCTCAGCTCAGCTCAGCTTCTTGAAGATCTGGCTTATCCAGAG 900
Qy 1540 CCACAGCGGTTGAGACATGAATTTGTACACAGATTGGCTTCGTGGTGGGCTCCCATACT 1599
Db 901 CCACAGCGAGTTGAGACATGAATTTGTACACAGCTGGCTTCGTGGTGGGCTCCCGTACT 960
Qy 1600 GAGGTGACACAAAGTGAAACCAAGCACTGTGGCGTCTCCAGAGCTGCTCGGAGTGTATC 1659
Db 961 GAGGTGACACAAAGTGAATAACCAACTGTGGCGTCTCCAGAGCTGCTCAGAGTGCATC 1020
Qy 1660 CTGGCCAGGACCCGCTGTGGCGCTGGAGCTTCCGGCTTGATGCTGTGTGGCCACAGCC 1719
Db 1021 CTGGCCAGGACCCAGTCTGTGCTCGAGCTTCCGGCTGGATGATGTGTGGCCCATGCC 1080
Qy 1720 GGCAGGACCGCGGGATGGTTCAAGATATAGATCAGCGGATGTCTTCTTTGTGTCCA 1779
Db 1081 GGGGAGCAGCGAGGGTTGGTCCAAGACATAGATCAGCAGATGTCTCTCTTTGTGTCT 1140
Qy 1780 AAAGAACCTGGAGAACATCCGTTAGTGTGAAAGTTCGGTGGCTACTGTGGGCCACAGTG 1839
Db 1141 AAAGAGCTTGAGAACGTCACAGTAGTGTGTTGAAGTTCGCGTGGCTACAGCTGCGCATGTG 1200
Qy 1840 GTCTGTGCATGTTCCCGCAGTCTCTGCTGGGCATCTGTGTGGCACCAAGCCAGTGGGA 1899
Db 1201 GTCTGTGCATGTTCTCCAGCTCAGCATGGGCATCTGTGTGGCACCAAGCCAGTGGGA 1260
Qy 1900 GTGACTGGCTCACTCCCGGAGGATGGACTAGAGTGTGTGTGACCCCAAGGGGCCATG 1959
Db 1261 GTGACTGCACTCAACCCCGCGGATGGACTGGAGTGTGTGTGACCCCAAGGGGCCATG 1320
Qy 1960 GGGGCTTATGCTTCGAGTGTCAAGAGGTGGAGCCCGCCGCTGGTGGCTGTATAGC 2019
Db 1321 GGGGCTTATGCTGTGAATGTCAAGAGGGTGGGACAGCCCATATGTTGGTACAGCTTACAGC 1380
Qy 2020 TTGCTGTGGGACAGCCAGCGGGACCTCAACCGGGCCACACCGTGTGGGGGCTTGGGA 2079
Db 1381 TTGATGTGGGACAGCCAGCAGATGTCTCCGAGCCGGGCCACAC- ---AGTGGGGCGGGA 1437
Qy 2080 TTGGTGTGCTTCTCTCGGTGTCTTTCAGCATCTCTCACTCTCTCTCTCTCTCTCTCTCT 2139
Db 1438 CTGGCTGGCTTCTCTTGGGATCTCGCAGCATCTCTGACTCTCTCTCTCTCTCTCTCTCT 1497
Qy 2140 CGTCAGCAGCTCGGCGACAGAGGAGCTTCTAGCTAGACAGAGGTGGGCTTATAGTCTG 2199
Db 1498 CGTCAGCAGCGACGCGACAGAGGAACTTCTGCTAGAGACAAAGGTGGGCTTGGAGCTG 1557
```

RESULT 8

AAV07282 .
ID AAV07282 standard; cDNA to mRNA; 1761 BP.

XX AAV07282;

XX AC AC

XX DT 08-SEP-1998 (first entry)

XX Human semaphorin W encoding cDNA.

XX Human; semaphorin W; nerve extension inhibitor; antiallergic; anticancer;

XX immunosuppressant; gene therapy; diagnosis; research reagent; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..1764

XX FT /*tag= a

XX FT /product= "semaphorin W"

XX WO9815628-A1.

XX PN

XX 16-APR-1998.

XX XX

XX PF 03-OCT-1997; 97WO-JP03549.

XX XX

XX PR 09-OCT-1996; 96JP-0287636.

XX XX (SUMU) SUMITOMO PHARM CO LTD.

XX PA Kikuchi K, Kimura T;

XX PI

XX XX

XX DR WPI; 1998-261015/23.

XX DR P-PSDB; AAW51314.

XX XX

XX PT Nerve extension inhibitor protein semaphorin W - is useful as

XX PT therapeutic drug and diagnostic and research reagent

XX XX

XX PS Claim 2; Page 67-68; 90pp; Japanese.

XX CC

XX The present sequence encodes human semaphorin W. Semaphorin W and

XX CC its derivatives are nerve extension inhibitors which are useful as

XX CC antiallergic, immunosuppressant and anticancer agents. The DNA

XX CC encoding semaphorin W can also be used in gene therapy, e.g. using

XX CC a viral vector. The proteins, peptides, DNA and antibodies which

XX CC recognise the protein or peptides, can be used as diagnostic or

XX CC research reagents. Semaphorin W can be used as a screen for

XX CC semaphorin W antagonists with possible therapeutic use.

XX Sequence 1761 BP; 349 A; 511 C; 521 G; 380 T; 0 other;

Query Match				34.7%;	Score 1391.2;	DB 19;	Length 1761;
Best Local Similarity				87.2%;	Pred. No. 1.8e-311;		
Matches 1538;				Conservative 0;	Mismatches 223;	Indels 3;	Gaps 1;
QY	640	GGGGGCGTCTCTACACGCCCACTGTGAAGAATCTTCCGGGACTGAGCCCATCTCC	699				
DB	1	GGGGGTGCTCTATGCTGCCACTGTGAATACTACCTGGGACGGAGCAATATACCC	60				
QY	700	CGAGTGTGGGTCGAGCTGAGGACTGGATTCCGAACAGAGACCTTGTCACTCTGGCTTAAT	759				
DB	61	AGAGCAGTGGGTCGTCGGAGACTGGATTCCGACAGATACCTTCCCTTCTGGCTGAAC	120				
QY	760	GCTCCAGCCTTTGTCGAGCTATAGTCTCGACCCAGCTGAGTGGGGGATGAAGATGA	819				
DB	121	GCCCCAGCCTTTGTCGACCGCTGGCTTGAAGCCAGCCGAATGGGGGATGAAGATGA	180				
QY	820	GACGATGAATCTTTTCTTACGGAGACCTCCCGAGTGTGGACTCTATAGGCG	879				
DB	181	GACGACGAATCTACTTCTTTTACGGAGACTTCCCGAGCATTTGACTCATACGAGCG	240				
QY	880	ATCAAGTCCCAAGAGTGGCCGAGTGTGTGGGGGACCTTGGGGGAGGAACCTT	939				
DB	241	ATTAAGTCCCAAGGCTGGCCGTGTGTGGGGGACCTCGGGGGCGGAGACCTC	300				
QY	940	CAGCAGAGATGACGACGTTTCTGAAGCTGACCTGTGTGCCAGGCCCCGAGCATGC	999				
DB	301	CAGCAGAGATGACGACGTTTGAAGCTGACCTGTCTGTCCAGGGCTGAGCATGC	360				
QY	1000	CGGGCTCCGGGTCTCGAGCTATGCGAGAGCTTCGGCTCAGCTTGGAGCGGAACC	1059				
DB	361	CGGGCTCCAGGTCTCGAGGATGTGTGTGTTCGACCTTGGAGCTTGGGGCAGGGACT	420				
QY	1060	CCATCTTTTATGGGATCTTTTCTCCAGTGGGAAGGAGTGCCATCTCTGCTGTGT	1119				
DB	421	CCATCTTTTATGGGATCTTTTCTCCAGTGGGAAGGAGTGCCATCTCTGCTGTGT	480				
QY	1120	GCTTCGACCCCAAGACATCGGGCAGTGTGTAATGTCCGTTTGAAGAGCTAAACAT	1179				
DB	481	GCTTCGACCCCAAGACATTCGGACAGTGTGTAATGTCCGTTTGAAGAGCTAAACAT	540				
QY	1180	GACTCAACAGGGGACTGCCTGTATGACACAGAGTGGCCCCAGCCAGACCTGGAG	1239				
DB	541	GACTCAACAGGGGACTGCCTGTATGACACAGAGTGGCCCCAGCCAGACCTGGAG	600				
QY	1240	TGATTCGCAACAAATGAAGTCCAGCAGTTTGGATCTCTACCTCCCTGCGACACCG	1299				
DB	601	TGATTCGCAACAAATGAAGTCCGCGACITTTGCTCTCTCTCCCTGCTTACCGC	660				
QY	1300	GTGCTCAGCTTTATCAGAGACACCTCTCATGAGACGGCCGTTTCCGGCTGACGG	1359				
DB	661	GTACTCAGCTTTATCCGGGACACCCACTCATGAGAGCGGAGTGTTCAGCTGATGG	720				
QY	1360	CGCCCCCTGCTGCTACAGATACGCTATCTCAGAGTGTGGGCCACAGGCTGACC	1419				
DB	721	CACCCCCCTGCTGCTACAGATACAGCTATCTCAGAGTGTGGGCCACAGGCTGACC	780				
QY	1420	AGCCTCTCAGGAAAGATATGACGTGCTACCTTGGGACAGAGGATGGACACCTCCAC	1479				
DB	781	AGCCTCTCAGGAAAGATATGATGCTCTACCTTGGGACAGAGGATGGACACCTCCAC	840				
QY	1480	CGGGCTGTGGGATGGAGCTCAGCTCAGTGTCTTGGAGGATCTGGCTTGTCCAGAA	1539				
DB	841	CGAGCAGTGTGGGATCGGAGCTCAGCTACGCTTCTTGAAGATCTGGCCCTTATCCACAG	900				
QY	1540	CCACAGCGGTTGAGAGCATGAATGTACACAGATTGCTCCTGGTGGCTCCCATACT	1599				
DB	901	CCACAGCGGTTGAGAGCATGAATGTACACAGATTGCTCCTGGTGGCTCCCGTACT	960				
QY	1600	GAGGTGACACAAGTGAACACACGACGATGTGGCCCTCTCCAGAGCTGTCTGAGTGTATC	1659				
DB	961	GAGGTGACACAAGTGAATACAAACCAACTGTGGCGCTCTCCAGAGCTGTCTCAGAGTGCATC	1020				
QY	1660	CTGGCCCCAGGCCGCTGGCCCTGGAGCTTCCGGCTTGATGTGTGTGGCCCCACGCC	1719				

RESULT 9
AA86128
ID AAX86128 standard; DNA; 799 BP.
XX
AC AAX86128;
XX
DT 15-SEP-1999 (first entry)
XX
DE EST sequence for DNA encoding SBSEPMN1.
XX
KW SBSEPMN1; semaphorin; neurodegeneration; spinal injury; neuropathy;
KW neuromuscular disorder; muscular dystrophy; psychiatric disorder;
KW inflammatory disorder; developmental malformation; EST;
KW expressed sequence tag; immune system disorder; cancer;
KW viral infection; vaccine; ss.
XX
OS Homo sapiens.
XX
PN EP933425-Al.
XX
PD 04-AUG-1999.

DB	1021	CTGGCCCCAGGCCAGTCTGTGCCGTGGAGCTTCCGGCTGGATGAGTGTGTGCCCATGCC	1080				
QY	1720	GGCGAGCACCGGGGATGTTCAAGATATAGAGTCAGCGGATGTCTCTTTTGTGTCCA	1779				
DB	1081	GGGGAGCACCGGAGGTGTGTCGAAGATAGAGTCAGAGATGTCTCTCTTTGTGTCT	1140				
QY	1780	AAAGAACCTGGAGAACATCCCGTAGTGTGAAGTTCCGGTGGCTACTGTGGGCCACGTC	1839				
DB	1141	AAAGAGCCTGGAGAACGTCCTAGTGTGAAGTTCCCGTGGCTACAGCTCGCATGTG	1200				
QY	1840	GTCTGTGCATGTTCCCGCAGTTCTCCCTGGGATCCTGTGTGTGACACAGCCAGTGA	1899				
DB	1201	GTCTTGCATGTTCTCAAGCTCAGCATGGGATCCTGTGTGTGACACAGCCAGTGA	1260				
QY	1900	GTGACTGCGCTCACTCCCGGGAGGATGAGCTAGAGTGGTGTACCCAGGGGCCATG	1959				
DB	1261	GTGACTGACCTACCCCCCGGGGATGAGCTGGAGTGGTGTGACCCAGGGGCCATG	1320				
QY	1960	GGGGCTTATGCTTGGAGTGTGAGGAGTGGAGCGCCGCGGTGGCTGTATAGC	2019				
DB	1321	GGCGCTTATGCTGTGAATGTGAGAGGTGGGCGAGCCCATGTGTAGCAGCTTACAGC	1380				
QY	2020	TTGGTGTGGGCGAGCCAGCGGGACCTCAAACCGGGGCCACACCGTTGTGGGGCTGA	2079				
DB	1381	TTGGTATGGGCGAGCCAGCGAGATGCTCCGAGCGCGGCCACAC	1437				
QY	2080	TTGGTGTGCTTCTCTGCTGTGCTTGTGACGATCCCTCACTCTCTCTGATTTGCTGC	2139				
DB	1438	CTGGCTGGCTTCTTCTTGGGATCTCGCAGATCCCTGACTCTCATTTGATTTGCTCG	1497				
QY	2140	CTCAGCAGCGTCGCGCAGAGAGGAGTCTTAGCTAGAGACAAGTGGGCTTAGATCTG	2199				
DB	1498	CTCAGCAGCGACGCGACAGAGGAACTTCTGGCTAGAGACAAGTGGGCTGGACCTG	1557				
QY	2200	GGGGCTCACCTTCTGGGACACAAAGCTATAGTAGAGACCTCCCTCTCTTTCGCTGAA	2259				
DB	1558	GGGGCTCACCTTCTGGGACACAAAGCTACAGCCAAAGCCCTCCCTCCCTCTCTGAA	1617				
QY	2260	GATGAACGGCTGCCCTGCGCTTGGGTAAGCGGGGAGTGGTTTGGTGGCTTCCCTCA	2319				
DB	1618	GATGAGCGGTTGCCCTGGCCCTGGCCAAAGAGGGGAGTGGCTTGTGGATTTCAACA	1677				
QY	2320	CCCTTCTGCTGGATTCTTGGCCAAAGCCAGCCACATCCGGCTCACTGGGGCGCTCTA	2379				
DB	1678	CCCTTCTGCTGTGATCTTGGCCAAAGCCAGCCACATTCGGCTAACTGGGCTCTCTA	1737				
QY	2380	GCCAGGTGTGAGACCTCCATC	2403				
DB	1738	GCCACATGTATGAACATCCATC	1761				

```
XX 30-SEP-1998; 98EP-0203287.
XX PF
XX PR
XX 30-JUL-1998; 98GB-0016676.
XX PR
XX 30-JAN-1998; 98EP-0300693.
XX PR
XX 28-JUL-1998; 98GB-0016423.
XX PR
XX PA
XX (SMIK ) SMITHKLINE BEECHAM PLC.
XX PI
XX Doe TR, Hayes PD, Michalovich D;
XX PR
XX WPI; 1999-407151/35.
XX DR
XX P-PSDB; AAY23875.
XX PR
XX New semaphorin family polypeptide useful for treating spinal injury
XX PT
XX and muscular dystrophy
XX PS
XX Disclosure; Page 16; 29pp; English.
XX PR
XX The present sequence represents an expressed sequence tag (EST)
XX sequence, from which DNA encoding SBSEMN1, a semaphorin family
XX polypeptide, is derived. The polypeptide may be used to screen
XX for agonists or antagonists including antibodies. Measurement
XX of the level of SBSEMN1 protein and detection of a mutation in
XX its polynucleotide may also be used to diagnose a disease or
XX condition or susceptibility to a disease related to altered
XX expression or activity of SBSEMN1. These diseases and
XX conditions include neurodegeneration, spinal injury, neuropathies,
XX neuromuscular disorders, muscular dystrophy, psychiatric disorders,
XX inflammatory disorders, developmental malformations, disorders of the
XX immune system, cancer and viral infections. The protein may also be
XX used to treat these diseases and conditions by administration as a
XX vaccine.
XX PR
XX Sequence 799 BP; 174 A; 237 C; 216 G; 171 T; 1 other;
XX PR
XX Query Match 16.3%; Score 654; DB 20; Length 799;
XX Best Local Similarity 88.6%; Pred. No. 3.7e-141;
XX Matches 708; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
XX PR
XX QY 1126 CGACCCCAACACATCCGGGACAGTCTGAATGGTCCCTTTAGAGAGCTAAACATGACTGC 1185
XX DB 1 CGACCCCAACACATTCGGACAGTCTGAATGGTCCCTTTAGAGAGCTAAACATGACTGC 60
XX PR
XX QY 1186 AACAGGGGACTGCTGTCATGGACAAGAGTGGCCCGACGCCAGACCTGGAGAGTGATC 1245
XX DB 61 AACAGAGAGTCTGCTGCTGGACAATGATGTGCCCGACGCCAGACCTGGAGAGTGATC 120
XX PR
XX QY 1246 GCCAACAAATGAAGTCCAGCAGTGTGGATCTCTCTCCCTGCCAGACCGCGTCTC 1305
XX DB 121 ACCAACAAATGAAGTCCCGGCACTTTGGCTCATCTCTCTCCCTGCCAGCGGTACTC 180
XX PR
XX QY 1306 ACCTTTATCAGACACACCTCTCATGACGAGCGCCGTGTCGGGCTGAGCGCGCC 1365
XX DB 181 ACCTTCATCCGGGACACCCACTCATGGACAGGCCAGTGTTCAGCTGATGGCCACCCC 240
XX PR
XX QY 1366 CTGCTGTCTACTACAGATACAGCCTATCTCAGAGTCTGTGCCACAGGGTGACAGCCTC 1425
XX DB 241 CTGCTGTCTACTACAGTACAGCCTATCTCAGAGTCTGTGCCACAGGGTGACAGCCTC 300
XX PR
XX QY 1426 TCAGGGGAAAGATATGACGTCTCTACCTGGGGACAGAGATGACACCTCCACCGGGCT 1485
XX DB 301 TCAGGGGAAAGATATGATGTGCTCTACCTGGGGACAGAGATGACACCTCCACCGAGCA 360
XX PR
XX QY 1486 GTGGCCATTTGGAGTCTAGCTAGTCTCTTGGAGATCTGGCTGTTCACAGAACACACAG 1545
XX DB 361 GTGGCCATTTGGAGTCTAGCTAGTCTCTTGGAGATCTGGCTGTTCACAGAACACACAG 420
XX PR
XX QY 1546 CCGGTTTCAGAGACGAATGTTACACAGATTGGCTCTGTGGGCTCCCTACTACTGAGGTG 1605
XX DB 421 CCAGTTGAGAACATGAATGTTACACAGCTGGCTCTGTGGGCTCCCTACTACTGAGGTG 480
XX PR
XX QY 1606 ACACAGTGAACACACGAACTGTGGCGGTCTCCAGAGCTGCTCGGAGTGTATCTGGCC 1665
```

```
DB 481 ACACAACTGAATACAACTGTCGCGCTCTCCAGAGCTGCTCAGAGTGCATCTGGCC 540
QY 1666 CAGGACCCCTGTGCGGCTCGGAGCTTCGGCTTGATGCTTGTGTGGCCCGCCGCGGAG 1725
DB 541 CAGGACCCAGTCTGTGCTCGGAGCTTCGGCTTGATGCTTGTGTGGCCCGCCGCGGAG 600
QY 1726 CACCGCGGATGGTTCAAGATATAGAGTCAAGAGTCTCTCTCTTTGTGTCCAAAAGAA 1785
DB 601 CACCGAGGGTGGTCCAAAGACATAGAGTCAAGAGTCTCTCTCTTTGTGTCTAGAGAG 660
QY 1786 CTTGGAGAACATCCCGTAGTGTGTTGAAGTTCGGTGGCTACTGTGGGCCAGCTGGTCTG 1845
DB 661 CTTGGAGAACATCCAGTAGTGTGTTGAAGTTCGGTGGCTACTGTGGGCCAGCTGGTCTT 720
QY 1846 CCATGTTCCCGGAGTCTGCTGGGCATCTGTGTGGCACCAGCCAGTGGAGTGAAT 1905
DB 721 CCATGTTCTCAAGCTCAGCATGGGCATCTGTGTGGCACCAGCCAGTGGAGTGAAT 780
QY 1906 GCGCTCACTCCCGGAGGG 1924
DB 781 TCACTTACCCCGCGGG 799
XX AC
XX ABT32087;
XX DT
XX 08-MAY-2003 (first entry)
XX DE
XX NOVX DNA sequence SEQ ID No 13.
XX KW
XX Anti-HIV; cytostatic; antiarteriosclerotic; antidiabetic; haemostatic;
XX hypotensive; neuroprotective; anorectic; nootropic; antidepressant;
XX immunosuppressive; antiparkinsonian; dermatological; antinfertility;
XX immunomodulator; antilipaeamic; NOVX; atherosclerosis; hypertension;
XX cancer; tuberculous sclerosis; Alzheimer's disease; Parkinson's disease;
XX multiple sclerosis; cachexia; anorexia; obesity; infertility;
XX atherosclerosis; hypertension; scleroderma; haemophilia; dyslipidaemia;
XX diabetes; autoimmune disease; immunodeficiency; AIDS; chromosome mapping;
XX graft-versus-host disease; tissue typing; pharmacogenomic; human; gene;
XX ds.
XX OS
XX Unidentified.
XX PN
XX WO2003004618-A2.
XX PD
XX 16-JAN-2003.
XX PF
XX 03-JUL-2002; 2002WO-US21360.
XX PR
XX 05-JUL-2001; 2001US-303046P.
XX PR
XX 11-JUL-2001; 2001US-304502P.
XX PR
XX 12-JUL-2001; 2001US-305011P.
XX PR
XX 17-JUL-2001; 2001US-306085P.
XX PR
XX 04-OCT-2001; 2001US-326981P.
XX PR
XX 28-FEB-2002; 2002US-360923P.
XX PR
XX 12-MAR-2002; 2002US-363636P.
XX PR
XX 16-APR-2002; 2002US-373063P.
XX PR
XX 02-JUL-2002; 2002US-0188246.
XX PA
XX (CURA-) CURAGEN CORP.
XX PI
XX Anderson DW, Boldog FL, Casman SJ, Edinger SR, Gerlach VL;
XX Gorman L, Li L, Malyankar UM, Patturajan M, Peyman JA, Shenoy SG;
XX Shimkets RA, Vernet CAM, Voss EZ;
XX PR
XX WPI; 2003-221591/21.
XX DR
XX P-PSDB; ABJ37119.
XX PR
XX New isolated NOVX polypeptides and polynucleotides, useful for
```


CC substitutions in the amino acids. The NOVX polypeptides are useful in
CC manufacturing a medicament for treating a syndrome associated with a
CC human disease, which is a pathology associated with the NOVX polypeptide.
CC The NOVX polypeptides, nucleic acids and antibodies are useful for
CC diagnosing, preventing or treating diseases such as atherosclerosis,
CC hypertension, cancer, tuberculous sclerosis, Alzheimer's disease,
CC Parkinson's disease, multiple sclerosis, cachexia, anorexia, obesity,
CC infertility, atherosclerosis, hypertension, scleroderma, haemophilia,
CC dyslipidaemias, diabetes, autoimmune disease, immunodeficiencies, AIDS,
CC or graft-versus-host disease. The nucleic acids are useful as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine, or pharmacogenomics. This polynucleotide sequence represents
CC the coding DNA of a NOVX protein of the invention.

XX
SQ Sequence 777 BP; 153 A; 230 C; 208 G; 186 T; 0 other;

Query Match 13.2%; Score 527.4; DB 25; Length 777;
Best Local Similarity 90.2%; Pred. No. 6.8e-112;
Matches 564; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 193 CGCGTCCCGCGCTCAGTCCCGAGAACCTCGCTGCCATCTCCGAGGCTGACTCCTATCTC 252
Db 7 CGCGTCCCGCGCTCGTGCAGAACCTCGCTGCCATCTCCGAGGCTGACTCCTATCTC 66
Qy 253 ACCCGGTTTGAGCGTCTATACGTACAAATPACTCTGCTCTCTTGTGGATCTCTGCTCC 312
Db 67 ACCCGGTTTGAGCGTCTATACGTACAAATPACTCTGCTCTCTTGTGGATCTCTGCTCC 126
Qy 313 CACACACTTTACGTCGCTGACGCGGATGACATCTTCCCTTTAACCCCTCCCTCTCTGGG 372
Db 127 CACACACTTTATGTTGGCGCCCGGACACCATCTTCGCTTTATCCCTGCGCTCTCAGGG 186
Qy 373 GAAAGACCCCGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 432
Db 187 GAGAGACCCCGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 246
Qy 433 AAAGCGAAGAAAGAGGAGGAGTGCACAAATTTTATCCAGATCTTCGCCATTTGCAATGCC 492
Db 247 AAAGCGAAGAAAGAGGAGGAGTGCACAAATTTTATCCAGATCTTCGCCATTTGCAATGCC 306
Qy 493 TCTCACCTCTCTCAGTGGCGACCTTCGCTTTTTCATCCGAGTCCGAGTCCGAGTCCGAGT 552
Db 307 TCTCACCTCTCTCAGTGGCGACCTTCGCTTTTTCATCCGAGTCCGAGTCCGAGTCCGAGT 366
Qy 553 TCCAGTTTCCAGCAGGTTGAAGACTTGAGAGCGCGCGGGGAAATGCTCTTTTGGAGCCA 612
Db 367 TCCAGTTTCCAGCAGGTTGAAGACTTGAGAGTGGCGGGGAAATGCTCTTTTGGAGCCA 426
Qy 613 GCTCAACGGTCAGCAGCTGTAATGGCTGGGGGCTCTCTACACCGCCACTGTGGAAGAAC 672
Db 427 GCTCAGGGGTCAGCAGCTGTAATGGCTGGGGGCTCTCTATGCTGCGCACTGTGAAAGAAC 486
Qy 673 TTCCTGGGAGTGAAGCCATCATCTCCGAGCTCTGGTCCGAGTCCGAGTCCGAGTCCGAGT 732
Db 487 TACCTGGGAGTGAAGCCATCATCTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGT 546
Qy 733 ACAGAGACCTTGTCTATCTGCTTAATGCTCCAGCCTTTTGTGCGAGCTATGCTCTGAGC 792
Db 547 ACAGATACCTTGTCTATCTGCTTAATGCTCCAGCCTTTTGTGCGAGCTATGCTCTGAGC 792
Qy 793 CCAGCTGAGTGGGGGATGAAGATG 817
Db 607 CCAGCCGAATGGGGGATGAAGATG 631

RESULT 12

ID ABL89791 standard; cDNA; 671 BP.

XX ABL89791;

AC ABL89791;

XX ABL89791;

DT 24-MAY-2002 (first entry)

XX

DE Human polynucleotide SEQ ID NO 353.
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein; gene; ss.
OS Homo sapiens.
XX WO200190304-A2.
PN 29-NOV-2001.
PD 18-MAY-2001; 2001WO-US16450.
PF 19-MAY-2000; 2000US-205515P.
PR (HUMA-) HUMAN GENOME SCI INC.
XX Birse CE, Rosen CA;
PI WPI: 2002-122018/16.
DR P-PSDB; ABB89382.
XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
XX disorders -
PS Claim 4; SEQ ID NO 353; 2081pp + Sequence Listing; English.
XX The invention relates to novel genes (ABL89449-ABL90853) and proteins
CC (ABB9040-ABB90444) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 671 BP; 134 A; 176 C; 214 G; 140 T; 7 other;

Query Match 12.6%; Score 505.2; DB 24; Length 671;
Best Local Similarity 86.6%; Pred. No. 8.8e-107;
Matches 582; Conservative 7; Mismatches 10; Indels 3; Gaps 3;
Qy 1387 GCCTATCTCAGAGTCGTGGCCACACAGGTCACAGCTCTCAGGAAAGATATGACGTG 1446
Db 1 GCCTATCTCAGAGTCGTGGCCACACAGGTCACAGCTCTCAGGAAAGATATGATGTG 59
Qy 1447 CTCTACCTGGGGACA-GAGGATGGACACCTCCACCGGGCTGTGGCATTTGAGCTCAGCT 1505
Db 60 CTCTACCTGGGGACATGAGGATGGACACCTCCACCGAGCACTGCGGAGCTCAGCT 119
Qy 1506 CAGTGTCTTGGAGATCTGGCCCTTGTTCACAGACACACAGCCGTTTGAGACATGAATTT 1565
Db 120 CAGCGTCTTGAAGATCTGGCCCTTATTTCCAG-AGCACAGCAGTGTGAGAACATGAATTT 178
Qy 1566 GTACACAGATTTGGCTCTCTGGTGGGCTCCCTACTAGGTCAGACAAAGTGAACACAGCAA 1625
Db 179 GTACACAGCTGGCTCTCTGGTGGGCTCCCTACTAGGTCAGACAAAGTGAACACAA 238

Db 956 ATATGGGGGGGACGACCCCTGCGAGAGAGAGTGGACCACTTCCTGAAGCGGGGCTGG 1015
Qy 977 TGTGCCCAGGCGCCGAGCATGGCGGGCTCCGGGGTTCTGCAGGCTATGCGAGAGCTTC 1036
Db 1016 CATGCTCTGCCCGAAGCTGCGAGCTTACTTCAACAGCTGCGAGGCGATG---CACACCC 1072
Qy 1037 GGCCCTCAGCTGAGCGGGGAACCCCACTTTTATGGGATCTTTTCTCCAGTGGGAAG 1096
Db 1073 TGCAGGACACCTCTGCGCACAACACCACTTCTTTGGGTTTTCAGAGCACAGTGGGGTG 1132
Qy 1097 GAGTGCATCTCTGCTGTGCTGCTCCGACCCCAAGACATCCGGGCAGTGTGAATG 1156
Db 1133 ACATGTACCTGTGGGCATCTGTGAGTACCAGTTGGAGAGATCCAGCGGGTGTGAGG 1192
Qy 1157 GTCCCTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCTGTCTATGGACAACGAGG 1216
Db 1193 GCCCTATAGAGTACCATGAGGAGCCCGAGAGTGGGACCGCTACACTGACCT---G 1249
Qy 1217 TGGCCCAAGCCAGACCTGGAGAGTGCAATCGCCCAACACATGAAGCTCCAGCAGTTGGAT 1276
Db 1250 TACCCAGCCCTCGCCCTGGCTGCTGCAATTAACAACCTGGCATCGCGCCACGGCTACACCA 1309
Qy 1277 CCTCAGCTCTCCCTGCCAGACCGGCTGCTACCTTATCAGAGACCACTCTCATGGACA 1336
Db 1310 GCTCCCTGGAGCTACCCGACAACATCTCACTTCGTCGAAGAGCACCCGCTGATGGAGG 1369
Qy 1337 GGCCCGTGTTCGGGCTGAGCGGCGCCCTGCTGCTCACTACAGATACAGCCTATCTCA 1396
Db 1370 AGCAGTGGGGCTCGGTGGAGCGCCCTGCTGCTGAAGAGGGCACCACTTCACCC 1429
Qy 1397 GAGTGTGGCCACAGGCTGACAGCCTCTCAGGGAAAGATATGAGCTGCTCTACCTGG 1456
Db 1430 ACCTGTGGCGGACCGGTTACAGGACTTGATGGAGCCACTATACAGTGTCTTCATTG 1489
Qy 1457 GGACAGAGATGGACACCTCCAGCGGCTGTCGCAATTGGAGCTCAGCTCAGTGTCTTG 1516
Db 1490 GCACAGAGACGGCTGCTCAAGGCTGTGAGCCTGGGGCCCTGGGTTACCTGATTG 1549
Qy 1517 AGGATCTGGCCTTGT---CCCAAGACACAGCGGTTGAGAGCATGAAATTTGTACCAG 1573
Db 1550 AGGAGCTGCAGCTGTTTGACCAGGAGCCCATGAGAAGCCTGGTCTATCTCAGAGCAAGA 1609
Qy 1574 ATTGGCTCCTGGTGGGCTCCCACTACTGAGTGCACAAAGTGAACACAGCAACTGTGGCC 1633
Db 1610 AGCTGCTCTTTGGCGGCTCCGCTCTCAGCTGGTGCAGCTGCCCGTGGCGGACTGCATAA 1669
Qy 1634 GTCTCCAGAGCTGCTCGAGTGTATCCTGGCCAGGACCCGCTGTGCGCCTGGAGCTTCC 1693
Db 1670 AGTATCGCTCTGTGCACTGTGCTCGCCGGGACCCCTATTGCGCCTGGAGCGTCA 1729
Qy 1694 GGCTTGATGTTGTGGCCACGCGCGGA 1724
Db 1730 ACACAGCGCTGTGTGGCGTGGGTGGCCA 1760

Search completed: September 25, 2003, 19:53:04
Job time : 991.19 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2003, 19:14:04 ; Search time 14345.7 Seconds
(without alignments)
11429.597 Million cell updates/sec

Title: US-09-284-180a-1
Perfect score: 4008
Sequence: 1 gccgagccgcgcagtagc.....aaaaaaaaaaaaaaaaaaaa 4008

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_or.*
21: em_ov.*
22: em_pat.*
23: em_ph.*
24: em_pi.*
25: em_ro.*
26: em_sts.*
27: em_un.*
28: em_vi.*
29: em_hgt_hum.*
30: em_hgt_inv.*
31: em_hgt_other.*
32: em_hgt_mus.*
33: em_hgt_pln.*
34: em_hgt_rod.*
35: em_hgt_mam.*
36: em_hgt_vrt.*
37: em_sy.*
38: em_hgtgo_hum.*
39: em_hgtgo_mus.*
40: em_hgtgo_other.*
41: em_hgtgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4008	100.0	4008	10	AB002563	AB002563 Rattus no
2	3109.2	77.6	4002	10	AB021291	AB021291 Mus muscu
3	2173	54.2	25079	2	AC135520	AC135520 Rattus no
4	2077.2	51.8	2354	10	AF038652	AF038652 Mus muscu
5	1954	48.8	4293	9	HS0801526	AL136552 Homo sapi
6	1875.2	46.8	2971	6	AX136143	AX136143 Sequence
7	1875.2	46.8	2971	6	BD123524	BD123524 Secretary
8	1875.2	46.8	2971	9	AK075384	AK075384 Homo sapi
9	1867.8	46.6	2948	9	BC038411	BC038411 Homo sapi
10	1673.6	41.8	2893	6	AX003081	AX003081 Sequence
11	1673.6	41.8	2893	6	BD177663	BD177663 SBSENN1 p
12	1673.6	41.8	2893	6	E35443	E35443 SBSENN1 p
13	1673.6	41.8	2894	6	AX003083	AX003083 Sequence
14	1673.6	41.8	2894	6	BD177664	BD177664 SBSENN1 p
15	1673.6	41.8	2894	6	E35444	E35444 SBSENN1 p
16	1483.8	37.0	162691	10	AB022311S6	AB022311 Mus muscu
17	1483.8	37.0	162691	10	AC003061	AC003061 Mouse Chr
18	1483.8	37.0	214837	10	AC007305	AC007305 Mus muscu
19	1475.8	36.8	109332	2	AC134899	AC134899 Mus muscu
20	1464	36.5	2672	9	AB021292	AB021292 Homo sapi
21	1201.4	30.0	2507	9	BC018361	BC018361 Homo sapi
22	1199.8	29.9	2400	9	AB022317	AB022317 Homo sapi
23	1034	25.8	1913	9	AF053369	AF053369 Homo sapi
24	692.8	17.3	2669	9	AB047604	AB047604 Macaca fa
25	654	16.3	799	6	AX003085	AX003085 Sequence
26	654	16.3	799	6	BD177665	BD177665 SBSENN1 p
27	654	16.3	799	6	E35445	E35445 SBSENN1 p
28	539	13.4	110000	2	AL451003_2	Continuation (3 of
29	539	13.4	126228	9	AC006543	AC006543 Homo sapi
30	539	13.4	234523	9	AC006544	AC006544 Homo sapi
31	527	13.1	163903	9	AC007387	AC007387 Homo sapi
32	477.2	11.9	64154	2	AL451003_3	Continuation (4 of
33	408	10.2	669	6	AX136484	AX136484 Sequence
34	408	10.2	669	6	BD123724	BD123724 Secretary
35	373.6	9.3	3293	6	AX174731	AX174731 Sequence
36	251.8	6.2	2155	6	AX060303	AX060303 Sequence
37	248.6	6.2	2156	6	AX060313	AX060313 Sequence
38	248.6	6.2	2284	6	AX060305	AX060305 Sequence
39	248.6	6.2	2646	6	AX704763	AX704763 Sequence
40	248.6	6.2	3776	6	BD171174	BD171174 Novel gen
41	248.6	6.2	3776	9	AB051526	AB051526 Homo sapi
42	247.8	6.2	2558	6	AX528277	AX528277 Sequence
43	247.8	6.2	3112	6	AX528279	AX528279 Sequence
44	247.8	6.2	3781	6	AX376386	AX376386 Sequence
45	247	6.2				

ALIGNMENTS

RESULT 1
AB002563
LOCUS
DEFINITION
Rattus norvegicus mRNA for semaphorin W, complete cds.
ACCESSION
AB002563
VERSION
AB002563.1 GI:4519426
KEYWORDS
Semaphorin W.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (sites)
AUTHORS
Encinas,J.A., Kikuchi,K., Chedotal,A., de Castro,F., Goodman,C.S.
and Kimura,T.

TITLE Cloning, expression, and genetic mapping of Sema W, a member of the
Semaphorin family
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (5), 2491-2496 (1999)
MEDLINE 99162633
PUBMED 10051670
REFERENCE 2 (bases 1 to 4008)
AUTHORS Kimura,T. and Kikuchi,K.
TITLE Direct Submission
JOURNAL Submitted (31-MAR-1997) Toru Kimura, Sumitomo Pharmaceuticals
Research Center, Discovery Research Laboratories II; 1-98, Kasugade
Naka 3-chome, Konohana-Ku, Osaka, Osaka 554, Japan
(E-mail:tkimura@sumitomopharm.co.jp, Tel.06-466-5228,
Fax:06-466-5491)

FEATURES

source

1..4008
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/sex="Female"
/tissue_type="Brain"
/dev_stage="Adult"
76..2406
/codon_start=1
/product="semaphorin W"
/protein_id="BAA73629.1"
/db_xref="GI:4519427"

CDS

/translation="MLARAERPRGPRPPPPPLSLLLALLSAPVCGRVPRS
VPTSLPTEADSYLTRFAASHTYNVSALLVDPASHLYVGARDSIFALTLPFSGERP
RIDMNVETHRONCKKKGKEDCHNEFTQILAIYNASHLLTCGTFADPKGCVIDVS
SFQVERLESGRGKCFPEPAQSAVMAAGVLIATVKNFLGTEPIIISRAVGAEDWI
RYETLSSMLNAPFAAMVLSPAEWDGDDDEIFFTETSRVLDSEVERIKVPRVAR
VCAGDLGRKTLQQRWTFLLKADLLCPGPEHGRASGVLAQMAELRPOPGAGTPIFYGI
FSSQEGRAISAVCAFRPQDIAVLNGPFRLEKHCNRLGPMVDNEVPQRPGEICIAN
NMKLOQSSLSLDRVLTFIRDHPLMDRPVPADGPLLVTTDTAYLRVVAHVTSLS
SKKEYDVLVLTGEDGHLHRAVRIGALSLEDLALFPEPQPVESMKLYHDLLVGSHT
EYVONTNCGRLQSCSEICLAQDPVCARSLDACVAHAGHRMWDIESADVSSL
CKPEGEHPVEFVAIVGVHVLVPCSPSSAWSCVHQPSPGVTAITPRRDGLEVVVT
PKMAGAYACEQEGAARVVAYISLVMSQSGPSNRHATVVGAGLVGLFLLGVLAASLT
LLIIGRRQRRORRELLARDKVLGLDGLGAPPSTTSYSQDPPSPSPEDERLPLALGRG
SGFGFPFPLDLDSPAHIRLTGAPLATCDETSI"

BASE COUNT 775 a 1135 c 1141 g 957 t

ORIGIN

Query Match 100.0%; Score 4008; DB 10; Length 4008;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4008; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCGAGGCCCGCGCAGTAGCGTACTAAGTAGAGGCTGCTGACGCGGCCCGCCACCGGCAC 60
DB 1 GCCGAGGCCCGCGCAGTAGCGTACTAAGTAGAGGCTGCTGACGCGGCCCGCCACCGGCAC 60

QY 61 CAGCGGGAGCCAGAGATGCTTCCAGGGCCGAGCGGCCCGCCCGGCCCGCCCGCGCCT 120
DB 61 CAGCGGGAGCCAGAGATGCTTCCAGGGCCGAGCGGCCCGCCCGGCCCGCCCGCGCCT 120

QY 121 CCGGTCTTTCCCTTCCCGCCCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
DB 121 CCGGTCTTTCCCTTCCCGCCCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180

QY 181 CCGGTGTGCGGCGCGCTCCCGCTCAGTGCCCAAGACCTGCTGCCATCTCCAGGCT 240
DB 181 CCGGTGTGCGGCGCGCTCCCGCTCAGTGCCCAAGACCTGCTGCCATCTCCAGGCT 240

QY 241 GACTCCTATCTCACCGGTTTTCAGCGTCTCATAGTACAATTAATCTGCTCTCCTCTGTG 300
DB 241 GACTCCTATCTCACCGGTTTTCAGCGTCTCATAGTACAATTAATCTGCTCTCCTCTGTG 300

QY 301 GATCCTGCTCCACACACTTTTACGTGCGTGCACGGGATAGCATCTTCGCTTTAACCCCTC 360
DB 301 GATCCTGCTCCACACACTTTTACGTGCGTGCACGGGATAGCATCTTCGCTTTAACCCCTC 360

QY 361 CCCTTCTCTGGGGAAGACCCCGAAGGATCGACTGGATGGTACCTGAGACTCACAGACAG 420

DB 361 CCCTTCTCTGGGGAAGACCCCGAAGSATCGATGGATGGTACTGAGACTCACAGACAG 420
QY 421 AACTGCAGGAAGAAAGGCAAGAAAGAGGAGCAATGTACAAATTTTATCCAGATTCTCGCC 480
DB 421 AACTGCAGGAAGAAAGGCAAGAAAGAGGAGCAATGTACAAATTTTATCCAGATTCTCGCC 480
QY 481 ATTGTCAATGCCTCTCACCTCCCTCAGTGCAGGACCTTCGCTTTTATCCGAAAGTCGGG 540
DB 481 ATTGTCAATGCCTCTCACCTCCCTCAGTGCAGGACCTTCGCTTTTATCCGAAAGTCGGG 540
QY 541 GTTATTGATGTGTCAGATTTCACAGAGGTTGAAGACTTGAGAGCGCGCGGGGAAATGT 600
DB 541 GTTATTGATGTGTCAGATTTCACAGAGGTTGAAGACTTGAGAGCGCGCGGGGAAATGT 600
QY 601 CTTTTTGAGCCAGCTCAACGGTCAGCAGCTGTAATGGCTGGGGCGCTCTCTACACCGCC 660
DB 601 CTTTTTGAGCCAGCTCAACGGTCAGCAGCTGTAATGGCTGGGGCGCTCTCTACACCGCC 660
QY 661 ACTGTGAAGAACTTCCTGGGGACTGAGCCCATCATCTCCGAGCTGTGGTTCGAGCTGAG 720
DB 661 ACTGTGAAGAACTTCCTGGGGACTGAGCCCATCATCTCCGAGCTGTGGTTCGAGCTGAG 720
QY 721 GACTGGATTTCGAACAGAGACCTTGTTCATCCTGCTTAATGCTCCAGCCTTTGTGCGAGCT 780
DB 721 GACTGGATTTCGAACAGAGACCTTGTTCATCCTGCTTAATGCTCCAGCCTTTGTGCGAGCT 780
QY 781 ATGGTCTCAGCCCCAGCTGAGTGGGGGATGAAGATGGAGAGCATCAAAATCTTTTTTTTC 840
DB 781 ATGGTCTCAGCCCCAGCTGAGTGGGGGATGAAGATGGAGAGCATCAAAATCTTTTTTTTC 840
QY 841 TTACAGGGACCTCCCGAGTGTGGACTCTATGAGCGCATCAAGTCCCAAGAGTGGCC 900
DB 841 TTACAGGGACCTCCCGAGTGTGGACTCTATGAGCGCATCAAGTCCCAAGAGTGGCC 900
QY 901 CGAGTGTGTCGGGGGACCTTTGGGGCAGAGAACCTTTCAGCAGAGATGGAGCAGCTTT 960
DB 901 CGAGTGTGTCGGGGGACCTTTGGGGCAGAGAACCTTTCAGCAGAGATGGAGCAGCTTT 960
QY 961 CTGAAGGCTGACCTGCTGTGCCAGGGCCGAGCATGGCGGCTTCCGGGGTTCGCGAG 1020
DB 961 CTGAAGGCTGACCTGCTGTGCCAGGGCCGAGCATGGCGGCTTCCGGGGTTCGCGAG 1020
QY 1021 GCTATGGCAGAGCTTCGGCCTCAGCTGGAGGGGAACCCCATCTTTTATGGGATCTTT 1080
DB 1021 GCTATGGCAGAGCTTCGGCCTCAGCTGGAGGGGAACCCCATCTTTTATGGGATCTTT 1080
QY 1081 TCCTCCAGTGGGAAGGAGCTGCCATCTCTGCTGTGTGCTTCCGACCCCAAGACATC 1140
DB 1081 TCCTCCAGTGGGAAGGAGCTGCCATCTCTGCTGTGTGCTTCCGACCCCAAGACATC 1140
QY 1141 CCGGCAGTGTCTGAATGGTCCCTTTAGAGAGCTTAAACATGACTGCAACAGGGGACTGCCT 1200
DB 1141 CCGGCAGTGTCTGAATGGTCCCTTTAGAGAGCTTAAACATGACTGCAACAGGGGACTGCCT 1200
QY 1201 GTCATGGAAACAGAGTGTGCCCCAGAGCTGGAGAGTGCATGCCCAACATGAAG 1260
DB 1201 GTCATGGAAACAGAGTGTGCCCCAGAGCTGGAGAGTGCATGCCCAACATGAAG 1260
QY 1261 CTCACAGAGTTTGGATCTCCTCCTCTGCTGCCAGACCGGGTGTCTACCTTTTATCAGAGAC 1320
DB 1261 CTCACAGAGTTTGGATCTCCTCCTCTGCTGCCAGACCGGGTGTCTACCTTTTATCAGAGAC 1320
QY 1321 CACCTCTCATGGACAGGCGCTGTTCGGGTGACGGCGCCCTGCTGCTGCTACTACA 1380
DB 1321 CACCTCTCATGGACAGGCGCTGTTCGGGTGACGGCGCCCTGCTGCTGCTACTACA 1380
QY 1381 GATACAGCCTATCTCAGAGTGTGCCCCACAGAGGTGACAGGCTCTCAGGGAAGAATAT 1440
DB 1381 GATACAGCCTATCTCAGAGTGTGCCCCACAGAGGTGACAGGCTCTCAGGGAAGAATAT 1440
QY 1441 GACGTGCTCTACCTGGGACAGAGGATGGACACCTCCACCGGCTGTGGCATTTGGAGCT 1500
DB 1441 GACGTGCTCTACCTGGGACAGAGGATGGACACCTCCACCGGCTGTGGCATTTGGAGCT 1500

Db	1441	GACGTGCTTACCTGGGACAGAGGATGGACACCTCCACCGGCTGTGCGCATGGAGCT	1500
Qy	1501	CAGCTCAGTCTCTGGAGGATCTGGCCCTGTCTCCAGAAACACAGCCGGTGGAGCATG	1560
Db	1501	CAGCTCAGTCTCTGGAGGATCTGGCCCTGTCTCCAGAAACACAGCCGGTGGAGCATG	1560
Qy	1561	AAATTGTACACGATGGCTTCCTGGTGGGTCCCATCTACTGAGGTGACACAAGTGAACACC	1620
Db	1561	AAATTGTACACGATGGCTTCCTGGTGGGTCCCATCTACTGAGGTGACACAAGTGAACACC	1620
Qy	1621	AGCAACTGTGGCGCTCTCCAGAGCTGCTCGAGATGTATCTGCCAGGACCCCGTGTGC	1680
Db	1621	AGCAACTGTGGCGCTCTCCAGAGCTGCTCGAGATGTATCTGCCAGGACCCCGTGTGC	1680
Qy	1681	GCCTGGAGCTTCGGCTTGATGCTTGTGGCCACAGCCGCGAGCACCCGGGATGTT	1740
Db	1681	GCCTGGAGCTTCGGCTTGATGCTTGTGGCCACAGCCGCGAGCACCCGGGATGTT	1740
Qy	1741	CAAGATATAGAGTCAGCGGATGTCTTCTTGTGTCCAAAGAACCTGGAGAATCC	1800
Db	1741	CAAGATATAGAGTCAGCGGATGTCTTCTTGTGTCCAAAGAACCTGGAGAATCC	1800
Qy	1801	GTAGTGTTCAGTTCGGTGGCTACTGTGGCCACAGTGTCTGCCATGTTCGCCAGT	1860
Db	1801	GTAGTGTTCAGTTCGGTGGCTACTGTGGCCACAGTGTCTGCCATGTTCGCCAGT	1860
Qy	1861	TCGCTCTGGGCATCTGTGTGGCACAGCCAGTGGAGTGACTGCGCTCACTCCCGG	1920
Db	1861	TCGCTCTGGGCATCTGTGTGGCACAGCCAGTGGAGTGACTGCGCTCACTCCCGG	1920
Qy	1921	AGGATGGATAGAGTGGTGTGACCCAGGGGCCATGGGGCTTATGTTCGGAGTGT	1980
Db	1921	AGGATGGATAGAGTGGTGTGACCCAGGGGCCATGGGGCTTATGTTCGGAGTGT	1980
Qy	1981	CAGGAGGTGAGCGCCCGCTGGTGGCTGTATAGCTTGGTGGGCGAGCAGCGG	2040
Db	1981	CAGGAGGTGAGCGCCCGCTGGTGGCTGTATAGCTTGGTGGGCGAGCAGCGG	2040
Qy	2041	GGACCTCAAAACCGGGCCACACCGTTGTGGGGCTGGATTGGTTCCTCGGT	2100
Db	2041	GGACCTCAAAACCGGGCCACACCGTTGTGGGGCTGGATTGGTTCCTCGGT	2100
Qy	2101	GTTCCTCAGCATCCCTCACTCTCCTGATTGGTGGCGCTCAGCAGCGTCGGCGACAG	2160
Db	2101	GTTCCTCAGCATCCCTCACTCTCCTGATTGGTGGCGCTCAGCAGCGTCGGCGACAG	2160
Qy	2161	AGGAGCTTCTAGCTACAGACAGGTGGCTTAGATCTGGGGCTCCACCTTCTGGGACC	2220
Db	2161	AGGAGCTTCTAGCTACAGACAGGTGGCTTAGATCTGGGGCTCCACCTTCTGGGACC	2220
Qy	2221	ACAAGCTATAGTCAGGACCTCCCTCTCCTTCGCTGAAGATGAACGGCTGCCCTGGCC	2280
Db	2221	ACAAGCTATAGTCAGGACCTCCCTCTCCTTCGCTGAAGATGAACGGCTGCCCTGGCC	2280
Qy	2281	CTGGGTAAAGGGGCGATGGTTTGGTGGCTTCCCTCCACCTTCCTGCTGGATTCCTGC	2340
Db	2281	CTGGGTAAAGGGGCGATGGTTTGGTGGCTTCCCTCCACCTTCCTGCTGGATTCCTGC	2340
Qy	2341	CCAGGCCACGCCACATCCGGCTCACTGGGCGCTCTAGCCAGCTGTATGAGACCTCC	2400
Db	2341	CCAGGCCACGCCACATCCGGCTCACTGGGCGCTCTAGCCAGCTGTATGAGACCTCC	2400
Qy	2401	ATCTAAAGCCGGGAAATGACTGCCAGCATGAGCAGTCTCTGGAATAGTGCTACCA	2460
Db	2401	ATCTAAAGCCGGGAAATGACTGCCAGCATGAGCAGTCTCTGGAATAGTGCTACCA	2460
Qy	2461	AGACCATGATCATGGCTGCTCTTCTTGGAGTCTGTGTTCACACATTAGTGTCTG	2520
Db	2461	AGACCATGATCATGGCTGCTCTTCTTGGAGTCTGTGTTCACACATTAGTGTCTG	2520
Qy	2521	TCCTCTGGACCTGGACCTGGCTTGGCCAGATTCCTGATTCATGAGATCAACCT	2580
Db	2521	TCCTCTGGACCTGGACCTGGCTTGGCCAGATTCCTGATTCATGAGATCAACCT	2580
Qy	2581	GTAACTTCTCGATGGCTCTTGTCTTGGGCCCATCAGCTTGTGGGTGAGTAAGGAC	2640
Db	2581	GTAACTTCTCGATGGCTCTTGTCTTGGGCCCATCAGCTTGTGGGTGAGTAAGGAC	2640
Qy	2641	ATAGGCCCGGAAAGGAATCAGTGTGGGTGAGTGTGGGGGTGTGGCCTCGCTT	2700
Db	2641	ATAGGCCCGGAAAGGAATCAGTGTGGGTGAGTGTGGGGGTGTGGCCTCGCTT	2700
Qy	2701	GTGTGGCTGTATGATTTCCAGTCTGCTGACTCTGGGAGCGCATGATCCCTGACTGC	2760
Db	2701	GTGTGGCTGTATGATTTCCAGTCTGCTGACTCTGGGAGCGCATGATCCCTGACTGC	2760
Qy	2761	CTTGAGATCTCTCCAACTCAGTTTCCCTTGTCTGGAAGAGTGTGTCTATACACTG	2820
Db	2761	CTTGAGATCTCTCCAACTCAGTTTCCCTTGTCTGGAAGAGTGTGTCTATACACTG	2820
Qy	2821	GTGTGGCTAGAGGCTGTCCATGTGTGATGAGACAGGCGGCTCGGTGCTT	2880
Db	2821	GTGTGGCTAGAGGCTGTCCATGTGTGATGAGACAGGCGGCTCGGTGCTT	2880
Qy	2881	TGGGAGTCTCGAGAGAGGTGGAAATGGGGAGCACTTAACCTCGGTAGCCAGTGAGG	2940
Db	2881	TGGGAGTCTCGAGAGAGGTGGAAATGGGGAGCACTTAACCTCGGTAGCCAGTGAGG	2940
Qy	2941	GAACACATCTCCCGTCCCATCACCCACAGCGCTTCTTAACTTTGAGCAAAAGTTCCC	3000
Db	2941	GAACACATCTCCCGTCCCATCACCCACAGCGCTTCTTAACTTTGAGCAAAAGTTCCC	3000
Qy	3001	AAAGTGACCTTCTGGGTGGGAGGAGAGACATGTGGCCCGCTCTCTCTCTCTCTCT	3060
Db	3001	AAAGTGACCTTCTGGGTGGGAGGAGAGACATGTGGCCCGCTCTCTCTCTCTCTCT	3060
Qy	3061	TTTCCCTTCTGGCTGCGCAACACCTGCGCTGCGCCAGTGTGGCTTCTCTCTCTCTCT	3120
Db	3061	TTTCCCTTCTGGCTGCGCAACACCTGCGCTGCGCCAGTGTGGCTTCTCTCTCTCTCT	3120
Qy	3121	AGGTGAGTCTCTCTCTCTCTCTCTTAAATGAACCTTCAACATTTCTAAATATT	3180
Db	3121	AGGTGAGTCTCTCTCTCTCTCTTAAATGAACCTTCAACATTTCTAAATATT	3180
Qy	3181	GGGGATGACAAATGACTTTTCCCGCAGAAAAGTGTGTAGGAAATACAAGCAGTTAAA	3240
Db	3181	GGGGATGACAAATGACTTTTCCCGCAGAAAAGTGTGTAGGAAATACAAGCAGTTAAA	3240
Qy	3241	GAAGATTGCTCAGTGTGCTTACCTTGGCTTAAAGCAGAGTCTCTCAGCTAGGCTC	3300
Db	3241	GAAGATTGCTCAGTGTGCTTACCTTGGCTTAAAGCAGAGTCTCTCAGCTAGGCTC	3300
Qy	3301	TGTGGACTCTCTGAAATTTGATGCTGTGGACTCTCTGAAATTTGATGCAAAAGTGTCT	3360
Db	3301	TGTGGACTCTCTGAAATTTGATGCTGTGGACTCTCTGAAATTTGATGCAAAAGTGTCT	3360
Qy	3361	GT	3420
Db	3361	GT	3420
Qy	3421	GT	3480
Db	3421	GT	3480
Qy	3481	TATAGTCAACACACTTGGGCGCATGGAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3540
Db	3481	TATAGTCAACACACTTGGGCGCATGGAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3540
Qy	3541	GGCTGT	3600
Db	3541	GGCTGT	3600
Qy	3601	CTTGGAAGCTAGGCTGTGGGCTGTGAAGGACACCATCTCTCTCTCTCTCTCTCTCTCT	3660
Db	3601	CTTGGAAGCTAGGCTGTGGGCTGTGAAGGACACCATCTCTCTCTCTCTCTCTCTCTCT	3660

D	b	844	TTCCTTACGAGACCCTCCCGAGCTCTGGACTCATATGAGCGCAATCAAGGTCCACGAGTG	903
Q	y	898	GCCCGAGTGTGTGCGGGGACCTTTGGGGGCAGGAAGAACCTTCAGCAGAGATGGAGCAGC	957
D	b	904	GCCCGTGTGTGCGGGGACCTTTGGGGCCGGAAGAACCTTCAGCAGAGATGGAGCAGC	963
Q	y	958	TTTTCTGAAGGCTGACCTGTGTGCCAGAGGCCCGAGCATGGCCGGCCTCCGGGGTTCTG	1017
D	b	964	TTTTCTAAGGCTGACCTGTGTGTCCAGGGCCCGAGCATGGAAGGGCCCTCGGGGGTTCTG	1023
Q	y	1018	CAGGCTATGCGCAGAGCTTCGGCCCTCAGCCCTGGAGCGGGAACCCCCATCTTTTATGGGATC	1077
D	b	1024	CAGGATATGACAGAGCTTCGACCTCAGCCTGGCGCGGGNACCCCCCTCTTTATGGCATC	1083
Q	y	1078	TTTTCTCCCACTGGGAAGAGGTGCCATCTCTGCTGTGTGTGCCCTTCGACCCCCAAGAC	1137
D	b	1084	TTTTCTCCCACTGGGAAGAGCGGCAATTCTGCTGTGTGTGCCCTTCGACCCCCAAGAC	1143
Q	y	1138	ATCCGGCAGTCTGAATGTGTCTTTTAGAGAGCTTAACACATGACTGCAACAGGGGACTG	1197
D	b	1144	ATCCGGCAGTGTGAATGTGTCTTTTAGAGAGCTTAACACATGACTGCAACAGGGGACTG	1203
Q	y	1198	CCTGTCTATGACAACAGAGTGCCTCACCTGCCAGACCTGGAGAGTGCATCGCCAAACACATG	1257
D	b	1204	CCTGTCTATGACAACAGAGTGCCTCACCTGCCAGACCTGGAGAGTGCATCGCCAAACACATG	1263
Q	y	1258	AAGCTCCAGCAGTTTGGATCCTCACCTGCCAGACCGCGTGTCTATATCAGA	1317
D	b	1264	AAGTTCAGCAGTTTGGATCCTCACCTGCCAGACCGTGTCTATATCAGA	1323
Q	y	1318	GACCACCTCTCATNGACAGCCCGTGTGCCGGCTGAGGGCCGCCCTGTGTGTCACT	1377
D	b	1324	GACCACCGCTCATNGACAGCCCGTGTGTCCAAGCTGATGGCCGCCCTGTGTGTCACT	1383
Q	y	1378	ACAGATACAGCTATCTCAGAGTGTGGGCCACAGAGTGACAGGCTGACAGGCTCTCAGGGAAGAA	1437
D	b	1384	ACGATACAGCTATCTCAGAGTGTGGCTTACCCGGGTGACAGGCTCTCAGGGAAGAA	1443
Q	y	1438	TATGAGTGTCTACCTGGGACAGAGATGGACACTCCACCGGCTGTGCCATTTGGA	1497
D	b	1444	TATGATGTCTACCTAGGACAGAGATGGACACTCCACCGGCTGTGCCATCGGA	1503
Q	y	1498	GCTCAGCTCAGTGTCTTGGAGATCTGGCCCTGTTCCCAGAACACAGCCGGTTGAGAC	1557
D	b	1504	GCTCAGCTCAGTGTCTTGGAGATCTGGCCCTGTTCCCAGAACACACCGTTGAGAC	1563
Q	y	1558	ATGAATTTGACACAGTTGGCTCCTGGTGGGCTCCCATCTCAGGTGACACAGTGAAC	1617
D	b	1564	ATGAATTTGACACAGTTGGCTCCTGGTGGGCTCCCATCTCAGGTGACACAGTGAAC	1623
Q	y	1618	ACCAGCACTGTGGCCGTCTCCAGAGCTGTCCGAGTGTATCTCGCCACGAGACCCCGTG	1677
D	b	1624	ACCAGCACTGTGGCCGTCTCCAGAGCTGTCCAGATGTATCTTGCCACGAGTCCCGTG	1683
Q	y	1678	TGCGCCTTGGAGCTTCCGGCTGTATGCTGTGTGGCCACCGCCGGCAGCACCCCGGATG	1737
D	b	1684	TGTGCTTGGAGCTTCCGGCTCGATGCTGTGTGGCCCATGTCAGCGCAGCACCGTGGGATG	1743
Q	y	1738	GTTCAAGATATAGAGTCAGCGGATGTCTCTTTGTGTCCAAAAGAACCTGGAGAACAT	1797
D	b	1744	GTTCAAGATATAGAGTCAGCGGATGTCTCTCTTTGTGTCCAAAAGAACCTGGAGAACAT	1803
Q	y	1798	CCGTAAGTGTGAAGTTCCGGTGGCTACTGTGGGCCAGCTGTCTGCTCCATGTTCCCCC	1857
D	b	1804	CCGTAAGTGTGAAGTTCCGTTGCTGTGGCTTACGGTGGGCCAGCTGTCTGCTCCATGTTCCCCC	1863
Q	y	1858	AGTTCGCTTGGGACCTCTGTGTGGCACCACCGCCAGCTGGAGTGTGCTGCTCACTCCC	1917
D	b	1864	AGTTCGCTTGGGACCTCTGTGTGGCACCACCGCCAGCTGGAGTGTGCTGCTCACTCCC	1923
Q	y	1918	CGAGGATGAGTATAGAGTGTGGTGTGACCCCCAGGGCCATGGGGCTTATGCTGCGAG	1977
D	b	1924	CGAAGGATGAGTATAGAGTGTGGTGTGACCCCCAGGGCCATGGGGCTTATGCTGCGAG	1983

Db	2280	CTGTGTGATCACAGCTAGTGTCTGTCTCTCGGACCTGGA	2319	Matches 3102; Conservative 0; Mismatches 870; Indels 316; Gaps 31;	
RESULT 5	HSMB01526	4293 bp	mRNA	linear	PRI 12-JUL-2002
LOCUS	DEFINITION	Homo sapiens mRNA; cDNA DKFZp761015121 (from clone DKFZp761015121); complete cds.			
ACCESSION	VERSION	AL136552			
KEYWORDS	SOURCE	AL136552.1	GI:13276610		
ORGANISM	Homo sapiens (human)				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 4293)				
TITLE	Poustka, A., Klein, M., Mewes, H.W., Gassenhuber, J. and Wiemann, S.				
JOURNAL	Direct Submission				
COMMENT	Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp761015121) is available at the R2PD in Berlin. Please contact the R2PD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@r2pd.de Further information about the clone and the sequencing project is available at http://mips.gsf.de/proj/cDNA/.				
FEATURES	Location/Qualifiers				
source	1..4293				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/chromosome="2p13"				
	/clone="DKFzp761015121"				
	/tissue_type="amygdala"				
gene	/clone_lib="761 (synonym: hamy2). Vector pSport1; host DH10B; sites NotI + SalI"				
	/dev_stage="adult"				
	1..4293				
	/gene="DKFzp761015121"				
CDS	112..2424				
	/gene="DKFzp761015121"				
	/note="semaphorin W"				
	/codon_start=1				
	/product="hypothetical protein"				
	/protein_id="CAB66487.1"				
	/db_xref="GI:13276611"				
	/db_xref="SWISS-PROT:O95754"				
	/translation="MPASARPRPGGQPTASPPFLLLAVLSGVPVGRVSRVPTS LPISADSLTRFAVPHYTYNYSLVLDVDPASHTLYVGARDTIFALSLPFSGERPRRIDW MVPEAHONCRKKKKEDECHNFVQILAIANASHLLTCGTFADPKCGVIDVSRFOQV PRLSEGRCKPEPAQSAAYMAGVLYAATVKNYLGTEPIITRAVGRADWIRDTL PSWLNAPFAAVALSPAEMGDDDEIYFFETTSRAFDSYERIKVPRVARCAGD LGGKRTLQORWTFLLKADLLCPGHEGRASSVLQDVAVLRELGAAGTPIFYGFSSOW EGATLSAYCAFQPDITRVINGPREFLKHGRLVVDNDVPOPRGECITNNMKLR HFGSLSLPDRVLVFIIRDPLMDRPFVADGPHLLVTTDYALRVVHRVTSLSGKEY DVLYLGTEDGHLHRAVRIGALSVELEDALEFPQPVENMKLYHSLWLVGSRTVFQV NTNNGRLQSCSECTLAQDPVCAISFRIDECVNAHAGEHRLVQDIESADVSLCPKEP GERPVFVPVATAHVVLPSPSASWQVHQPSGTALTTPRDGLEVVTTPGAMG AYACBCQGGAAHVVAAYSLVWGSORDAPSAHRAHTVAGLAGFFGLIAASLTLLIGR RQORRQRELLARQKVLGDLGAPPSGTTYSQDPPSPSPEDERLPLALAKRSGSGFGF SPFFLLDPCPSAHLRTGAPLATCDETSI"				
	4251..4256				
polyA_signal	/gene="DKFzp761015121"				
polyA_site	4270				
BASE COUNT	905	a	1171	c	1177
ORIGIN	g 1040 t				
Query Match	48.8%; Score 1954; DB 9; Length 4293;				
Best Local Similarity	72.3%; Pred. No. 0;				

1081 TCCTCCAGTGGGAAGGAGCTGCCATCTCTGCTGTGTGCTTCCGACCCCAAGACATC 1140
1102 TCTTCCAGTGGGAGGGCTACTATCTCTGCTGTGTGCTTCCGACCCCAAGACATC 1161
1141 CGGCGAGTGTGAATGGTCCCTTTAGAGAGCTAAACATGACTGCACAGGGGACTGCCT 1200
1162 CGGAGTGTGAATGGTCCCTTTCAGAGAACTAAACATGACTGCACAGAGGAGTGCCT 1221
1201 GTCATGACACAGAGGTGCCACGCCAGACTGGAGAGTGCATCGCCCAACAAATGAAG 1260
1222 GTCGTGACAAATGATGTGCCCCAGCCAGACTGGAGAGTGCATCACCAACAAATGAAG 1281
1261 CTCAGCAGTTTGGATCTCACTCTCCCTGCCAGACCGCGTGCACCTTTATCAGAGAC 1320
1282 CTCGGCAGCTTTGGCTCATCTCTCCCTGCTGACCGGTACTCACCTTCATCCGGGAC 1341
1321 CACCTCTCATGAGACGGCCGCTGTCCCGCTGACGGCCGCCCTGCTGGTCACTACA 1380
1342 CACCCACTCATGAGACGGCCAGTGTTCAGAGTGCATGGCCACCCCTGCTGGTCACTACA 1401
1381 GATACAGCTATCTCAGAGTGGTGGCCACAGGTGACACCGCTCTCAGGGAAAGAAATAT 1440
1402 GATACGCTATCTCAGAGTGGTGGCCACAGGTGACACCGCTCTCAGGGAAAGAGTAT 1461
1441 GACGTGCTATCTACCTGGGACAGAGATGGACACCTCCACCGGGCTGTGCGCAATGGAGCT 1500
1462 GATGTGCTCTACCTGGGACAGAGATGGACACCTCCACCGAGCAGTGGCGATCGGAGCT 1521
1501 CAGCTCAGTCTTGGAGGATCTGGCTTGTTCACAGACCTCCAGACCCAGCGGTTGAGAGATG 1560
1522 CAGCTCAGCTCTTGAAGATCTGGCTTATATCCAGAGCCACAGCCAGTGTGAAGACATG 1581
1561 AAATGTACACAGATTTGGCTCCTGTGGGTCCCATACTAGGTGACACAAAGTGAACACC 1620
1582 AAATGTACACAGTGGCTCCTGTGGTCCCGTACTGAGGTGACACAAAGTGAATACA 1641
1621 AGCAACTGTGGCTCTCCAGAGTGTCTGAGGTGTATCTGTCGACAGGACCCGCTGTGC 1680
1642 ACCAACTGTGGCTCTCCAGAGTGTCTGAGGTGTATCTGTCGACAGGACCCAGTCTGT 1701
1681 GCCTGGAGCTTCCGGCTTGATGCTGTGTGTCGCCACGCCGGCGAGACCCGGGATGGTT 1740
1702 GCCTGGAGCTTCCGGCTTGATGCTGTGTGTCGCCACGCCGGGAGGCTTGGTC 1761
1741 CAAGATATAGAGTCAAGGATGTCTCTTCTGTGTGTCCTCAAAAGAACCTGGAGAACATCCC 1800
1762 CAAGACATAGAGTCAAGAGTGTCTCTCTTGTGTCTTAAAGAGCCTGGAGAACGTCCA 1821
1801 GTAGTGTGAAGTTCGGTGGCTTACTGTGTGGCCACGCTGTGTCCTGCCATGTTCGCCAGT 1860
1822 GTAGTGTGAAGTTCGGTGGCTTACTGTGTGGCCACGCTGTGTCCTTGGCTTCTCCAAGC 1881
1861 TCTGCCCTGGACATCTGTGTGTCGACCCAGCCAGTGGAGTGTGCTGCTCACTCCCGGG 1920
1882 TCAGCATGGGATCCTGTGTGTGACACAGCCAGTGGAGTGTGCTGCTCACTCCCGGG 1941
1921 AGGATGACCTPAGAGTGGTGGTACCCAGCGGCTATGGGGCTTATGCTTGGAGTGT 1980
1942 CGGGATGACCTGGAGTGGTGGTGAACCCAGGGGCTATGGCGCTTATGCTGTGAATGT 2001
1981 CAGGAGGTGAGCGCCCGCTGGTGGCTGTATAGCTTGGTGTGGGCGAGCCAGCG 2040
2002 CAGGAGGTGGGCGAGCCATGTGTAGCAGCTTACAGCTTGGTATGGGCGAGCCAGCGA 2061
2041 GGACCTCAAAACCGGCCCCACACCGTGTGGGGCTGGATGGTGTCTTCTCCGGGT 2100
2062 GATGCTCGAGCCGGGGCCACAC---AGTGGGGGGGAGTGGTGGCTTCTTCTGGGG 2118
2101 GTTCTTGACAGTCCCTCAGTCTCTCTGATTTGGTGGCGCTCAGACGCTGCGGAGAC 2160
2119 ATTCTCGAGCATCCCTGACTCTCATTTCTGATTGGTGGCGCTGACGCGGCGAGAC 2178
2161 AGGAGCTTCTAGCTAGACAAAGTGGGCTTAGATCTGGGGCTCCACCTTCTGGGACC 2220
2179 AGGAACTTCTGGCTAGAGACAAGTGGGCTTGACCTTGGGGCTCCACCTTCTGGGACC 2238
2221 ACAAGCTATAGTCAAGACCTTCCCTCTCTCGCTTGGCTGAAGATGAACGGCTGCCCTGGCC 2280
2239 ACAAGCTACAGCCCAAGACCTTCCCTCCCTCTCTCTGAAGATGAGCGGTTCGCCCTGGCC 2298
2281 CTGGGTAAAGCGGGCAGTGGTGGTGGCTTCCCTCCACCTTCTCTCGCTGGATCTCTTGGC 2340
2299 CTGGCCAAAGAGGGCAGTGGCTTGGTGGATCTCAACACCTTCTCTCGCTTGAATCTTGGC 2358
2341 CCAAGCCAGCCACATCCGGCTCACTGGGGCTCTTAGCCAGTGTGATGAGACCTTCC 2400
2359 CCAAGCCAGCCACATTCGGCTAACTGGGGCTCTCTAGCCACATCTGATGAACATCC 2418
2401 ATCTAAAGCCGGGAAATGACTGCCACCATGAGCAGTCTCTGGAAGTCTGAGTGG----- 2454
2419 ATCTAGAGCTGGGCAATGACCATAGTGTATAAGTGAATGATCACTGGAACGGAGTCACTACT 2478
2455 -----CTACCAAGACCATGATCATGGCTGCTTCTTCTCTTGGAGTCT 2496
2479 GAGATGCTGGGGTCACTGGGCTTGAAGACCATCCAGCCTCTGAGTCTCTTGGAGTCT 2538
2497 TGTGTGTTTAC-----ACATTAAGTGTCTGTCTC-----TGACCTGGAACCTTGG 2540
2539 TGAGTGAATTAATGGATTTTATCTCTCTGAGCCTGGATGGCTTGGGGCCAG 2598
2541 CCTTTGCCAGATTCCTGATTTCTCATGAGAGATCAACCTTG----- 2581
2599 ACCTTTGCCATGATTCCTGATTTCCATGAGAAATCAGAACTGCTTCTTCGACGAAATCAGG 2658
2582 -----TAACTTCTGCGATGGCTCTTCTTGGGGCC 2614
2659 GCTTCCCTTAACTCTGAATCTCTGTAAACCTTCACTCCCTGGCCCCATCTTGGGGCC 2718
2615 ATCAGCTGTGGGTGGAGTAAAGACATAGGCCCGGAAAGGAATAGTGTGAGAGTAG 2674
2719 ATTAGTTTGGGATGGGACAGGACATATGACTTTGCTTCTGTTGGAGCCTG 2778
2675 -----TTGGGGCTGTGCTGCTGCTGCTTGTGGTGG 2707
2779 GCCGAAAGGAAGACCTTGGAGTGGTTGGGGGCAATGTGCCCTGATCTCTTGGGGTGG 2838
2708 CTGTATGATTTCCAGTCTGCTGA-----CTCTGGGAGCGCATGATCCCC----- 2753
2839 TTCTGCTTATTTCTCAAGTTTATCTGAATCTGTGGGAGTGCATGATCCCATGTTGCAA 2898
2754 -----TGACTGCTTGAGATCTCTCCAACTCAGTTTCCCCTTGTCTTGGAGAGTGT 2806
2899 TATGAGTCTCTGCCCTGAGATCTTCCCATCTCAGTTTTC-----TTCCATGAAGAGTAC 2956
2807 GTGTCTATACACTGGTGGCTTAGAGCCTTCCATGTGTCATGGAGACAGCGCCG 2866
2957 GTGTAATATACA-TAGTGTTCATAGAGACCTTGGCCACATGTGCATGAATGACTGGGCCGA 3015
2867 TGCCTCGTGTCTTT-----GGGGAGTCCGAGAGAAAGTTTGAATGGGG----- 2912
3016 TGCCTTAGAATATTTATGAGGATGGGAGGAGACAAATTGAATGGGGCTTCCCTGTC 3075
2913 ACAACTTAACCTCGGTAGCCAGTGAAG-----GAAAC 2945
3076 AGAACTTTAGACCTCTAGCCATAGAGGAATGGGCTTCTTCTTATAAAGGAACCTGAAAC 3135
2946 CACATGCCGCTCCCATCACCCACAGCGCTTCTTAACTTTG-----AGCAAGTTCCTCC 3000
3136 ACTGCCCTCTCCCATTTGCCCAACACCTTACTCTAAATTTGAGTAGAGAAAGCTCCT 3195
3001 AAAGTGAATCTTGGTGGGAGGAGCAGCAGAGATGTGGCCCGCTCTCTCTCTCTCTCT 3060
3196 AAAGTGAATCTTCCGGTGGGAGGAGCAGCAGTACAT-TACCCCATCTCTTCTTCTTGT 3254
3061 TTTCCCTTCTGGCTGCCAAACCACTGCCGTGCCCGCTTCTTCCCTGCTTGGAGTGG 3120

[illegible]

373	CCCTTCTCAGGGGAGACCCCGCAGGATTGACTGGATGGTTCTTGAGGCTCAGACAG	432	
421	AACTGCAGGAAGAACGAAAGAGAGACAAATGTTCACAAATTTATCCAGATCTCGCC	480	
433	AACTGTAGGAAGAACGAAAGAGAGACAAATGTTCACAAATTTGTCAGATCTCGCC	492	
481	ATTGTCAATGCCCTCTACCTCTCAGCTGGCGACCTTCGCTTTTGATCCGAAGTCGGG	540	
493	ATTGCCAATGCCCTCTACCTCTCAGCTTTGGCACCCTTCGCTTTTGATCCGAAGTCGGG	552	
541	GTATTATGATGTCTCCAGTTTTCAGCAGAGTTCGAAGACTTGAGAGCGCGCGGGGAAATGT	600	
553	GTATTATGATGTCTCCAGTTTTCAGCAGAGTTCGAAGACTTGAGAGTTCGAAGTTCG	612	
601	CCCTTTGAGCCAGCTCAACGGTTCAGCAGCTGTAATGGCTGGGGCGTCCCTCAACAGCC	660	
613	CCCTTTGAGCCAGCTCAGCGGTTCAGCAGCTGTAATGGCTGGGGGGTCCCTATGCTGCC	672	
661	ACTGTGAAGAACTTCTTGGGACTGAGCCCATCATCTCCCGAGCTGTGGTTCGAGCTGAG	720	
673	ACTGTGAAGAACTTCTTGGGACGAGCAATATCACAGAGCAGTGGGTCTGTCGCGAG	732	
721	GACTGGATTGCAACAGAGACCTTGTATCTCTGGCTTAATGCTCCAGCCTTTGTCGAGCT	780	
733	GACTGGATTGCGAGAGATACCTTGGCTTCTCTGGCTGAACGCCCGCAGCCTTTGTGCGAGCC	792	
781	ATGGTCTGAGCCAGCTGAGTGGGGGATGAAGATGGAGACGATGAATCTTTTTC	840	
793	GTGGCTTGAAGCCAGCCGAATGGGGGATGAAGATGGAGACGAGCAATCTACTTCTTC	852	
841	TTACAGGAGACTCCCGAGTGTGGACTCTTATGAGCGCATCAAGAGTCCCAAGAGTGGCC	900	
853	TTTACGAGACTTCCCGAGCTTCTGACTCATCAGAGCGCAATAAAGTCCCGAGTGGCC	912	
901	CGAGTGTGTGGGGGACCTTGGGGGAGGAAGACCTTTCAGCAGAGATGAGCAGACGTTT	960	
913	CGTGTGTGTGGGGGACCTCGGGGGCGGAAGACCTTCCAGCAGAGATGAGCAGACGTTT	972	
961	CTGAAGGCTGACTGCTGTGCCAGGCGCCAGAGATGGCGGGCTCCGGGGTCTGCAG	1020	
973	TTGAAGCTGACTGCTGTGTCCAGGGCTTGAGCATGGCGGGCTCCAGTGTCTGCAG	1032	
1021	GCTATGGCAGAGCTTCGGCCCTCAGCCTGGAGCGGAACCCCATCTTTTATGGATCTTT	1080	
1033	GATGTGTGTGCTTCGACCTGAGCTTGGGCGAGGACTCCCATCTTTTATGGCATCTTT	1092	
1081	TCTTCCAGTGGGAGGAGCTGCCATCTCTGCTGTGTGTGCTTCGAGCCCAAGACATC	1140	
1093	TCTTCCAGTGGGAGGAGCTTACTATCTCTGCTGTGTGCTTCGAGCCCAAGACATTT	1152	
1141	CGGGCAGTGTGAATGGTCCCTTTAGAGAGCTTAAACATGACTGCAACAGGGGACTGCCT	1200	
1153	CGGACAGTGTGAATGGTCCCTTTAGAGAGCTTAAACATGACTGCAACAGGAGTGCCT	1212	
1201	GTCATGGACACAGAGTGGCCCGAGCCAGACCTGGAGAGTGCATCGCAACAAACATGAAG	1260	
1213	GTCGTGGACAAATGATGTGCCCCAGCCAGAGCTGGAGAGTGCATCACCACAAACATGAAG	1272	
1261	CTCCAGCAGTTTGGATCTCATCTCTCTGCGCAGACCGCGTGCATCTTATCAGAGAC	1320	
1273	CTCCGCACTTTGGCTCATCTCTCTCTGCTGACGCGCGCGCTGCTGCTGCTGCTGCTGCT	1332	
1321	CACCTCTCATGAGAGCGCGCTTCCGGCTGACGCGCGCGCGCTGCTGCTGCTGCTGCTGCT	1380	
1333	CACCCACTCATGAGAGCGCGCTTCCAGCTGATGGCCACCGCGCTGCTGCTGCTGCTGCT	1392	
1381	GATACAGCCTATCTCAGAGTCTGTGGCCCAAGGCTGACAGCCTCTCAGGGAAGAAATAT	1440	
1393	GATACAGCCTATCTCAGAGTCTGTGGCCCAAGGCTGACAGCCTCTCAGGGAAGAGTAT	1452	
1441	GAGCTGCTTACTCTGGGACAGAGATGGACACCTCCACCGGGCTGTGCGCATTTGGAGCT	1500	
1453	GATGTGCTTACTCTGGGACAGAGATGGACACCTCCACCGGAGCTGCGGATCGGAGCT	1512	
1501	CAGCTCAGTGTCTTGGAGGATCTGGCCTTGTTCAGAGACACAGCCGTTTGAGAGCATG	1560	
1513	CAGCTCAGCGTCTTCTTGAAGATCTGGCCTTATTCAGAGACACAGCCAGTTGAGAACATG	1572	
1561	AAATTGTACACAGATTGGCTCCTGGTGGGCTCCCATACTGAGGTGACACAGTGAACACC	1620	
1573	AAATTGTACACAGTGGCTCCTGGTGGGCTCCCATCTACTGAGGTGACACAGTGAATACA	1632	
1621	AGCAACTGTGGCCGCTCTCCAGAGCTGCTCGAGGTATCTCTGGCCAGGACCCCGTGTGC	1680	
1633	AGCAACTGTGGCCGCTCTCCAGAGCTGCTCAGAGTGCATCTCTGGCCAGGACCCAGTCTGT	1692	
1681	GCCTGTGAGCTTCCGGCTTGATGCTGTGTGGCCACCGCGCGAGCAGCCCGGAGTGGTT	1740	
1693	GCCTGTGAGCTTCCGGCTTGATGCTGTGTGGCCACCGCGCGAGCAGCCAGGTTGGTC	1752	
1741	CAAGATATAGAGTCAGCGGATGCTCTCTCTTCTTGTCTCCAAAGAACCTGGAGAACATCCC	1800	
1753	CAAGACATAGAGTCAGCAGATGCTCTCTCTTCTTGTCTTAAAGAGCCTGGAGAACATCCC	1812	
1801	GTAGTGTTTGAAGTTCGGTGGCTACTGTGGGCCACGTTGCTCTGCCATGTTCCCCAGT	1860	
1813	GTAGTGTTTGAAGTTCGGTGGCTACAGCTCGCATGTGGTCTTTGCCATGTTCTCCAAGC	1872	
1861	TCTGCTGGGCATCCTGTGTGTGCCACAGCCCCAGTGGAGTGTACTGCGCTCACTCCCCGG	1920	
1873	TCAGCATGGCATCCTGTGTGTGCCACAGCCCCAGTGGAGTGTACTGCATCAACCCCCCG	1932	
1921	AGGATGAGCTAGAGTGTGTGTACCCAGGGCCATGGGGGCTTATGCTTGGCAGTGT	1980	
1933	CGGGATGGAAGTGTGGTGACCCAGGGCCATGGGCGCTTATGCTGTGAAATGT	1992	
1981	CAGGAGGTGGAGCCCGCGGTGGTGGCTGCTTATAGTCTTGTGTGGGCGAGCCAGCGG	2040	
1993	CAGGAGGTGGGCGAGCCCATGTGTGTAGCAGCTTACAGCTTGGTATGGGCGAGCCAGCGA	2052	
2041	GGACCTCTAAACCGGGCCACACCGTTGTGGGGGCTGGATGGTTGGCTTCTCTCTGGGT	2100	
2053	GATGCTCCGAGCCGGGCCACAC---AGTGGGGCGGAGTGGCTGGCTTCTTCTTGGGG	2109	
2101	GTTCTTGCAGATCCCTCAGTCTCTCTCTGATTGGTGCCTGACAGAGCTGCGGCGACAG	2160	
2110	ATTTCTCGCAGCATCCCTGACTCTCATTTGTTGGTGGCGCTCAGCAGCGCGAGCAG	2169	
2161	AGGAGCTTCTAGCTAGAGACAAGTGGGCTTTAGATCTGGGGCTTCACACCTTCTGGGACC	2220	
2170	AGGAACTTCTGGCTAGAGACAAGTGGGCTTGGACCTTGGGGGCTCCACCTTCTGGGACC	2229	
2221	ACAAGCTATAGTCAGGACCCCTCCCTCTCTTCCGCTGAAGATGAACGGCTGCCCTGGCC	2280	
2230	ACAAGCTACAGCAAGACCCCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2289	
2281	CTGGGTAAAGCGGGCAGTGGTTTGGTGGCTTCCCTCCACCCCTCTCTCTCTCTCTCTCTCT	2340	
2290	CTGGCAAGAGGGGAGTGGCTTGGTGGATTTCTCACCACCCCTCTCTCTCTCTCTCTCTCT	2349	
2341	CCAAGCCAGCCACATCCGGCTCAGTGGGGGCTCTAGCCAGCTGTGTGATGAGACCTCC	2400	
2350	CCAAGCCAGCCACATTCGGCTAACTGGGGCTCTCTTAGCCACATGTGTGAGAACATCC	2409	
2401	ATCTAAAGCGGGGAAAATGACTGCCAGCATGAGCAGTCTCTCTGGAAC	2448	
2410	ATCTAGAGCTGGGCAATGACCATGATGTATAGTGATCACTGGAAC	2457	
RESULT 7			
BD123524			
LOCUS	2971 bp DNA		
DEFINITION	Secretory protein or membrane protein.		
ACCESSION	BD123524		
VERSION	BD123524.1		
KEYWORDS	JP 2002017376-A/33.		
PAT 18-SEP-2002			

[illegible]

Db 1633 ACCAACTGTGGCCCTCCACAGAGCTGCTCAGAGTGCAATCTCGGCCACAGACCCAGTCTGT 1692
Qy 1681 GCCTGGAGCTTCCGGCTTGATGCTTGTGTGCCACCGCGGAGACACCGCGGATGTT 1740
Db 1693 GCCTGGAGCTTCCGGCTGGAGTGCTGTGCCATGCCGGGAGACCCGAGGTTGGTC 1752
Qy 1741 CAAGATATAGTACAGCGGATGCTCTCTTTTGTGTCCAAAGAACCTCGGAGAACATCCC 1800
Db 1753 CAAGACATAGATCAGCAGATGCTCTCTTTTGTGTCTTAAAGAGCCTGGAGAACGTCCA 1812
Qy 1801 GTAGTGTTTGAAGTTCGGTGGCTACTGTGGCCACAGTGGTCTGCTGCTTCCCTCCAGT 1860
Db 1813 GTAGTGTTTGAAGTTCGGTGGCTACTGTGGCCACAGTGGTCTGCTGCTTCCCAAGC 1872
Qy 1861 TCTGCCCTGGGCATCTGTGTGGCACACAGCCAGTGGAGTACTGGGCTCACTCCCGCG 1920
Db 1873 TCAGCATGGGCATCTGTGTGGCACACAGCCAGTGGAGTACTGCACTCACCCCGCG 1932
Qy 1921 AGGATGGACTAGAGTGGTGGTGACCCAGGGGCCATGGGGGCTTATGCTTGGAGTGT 1980
Db 1933 CGGATGGACTGGAGTGGTGGTGACCCAGGGGCCATGGCGCTTATGCTGTGAATGT 1992
Qy 1981 CAGGAGGTGGAGCCCGCGCTGGTGGCTCTTATAGCTTGGTGTGGGGCAGCAGCGG 2040
Db 1993 CAGGAGGTGGGGCAGGCCATGTGGTAGCAGCTTACAGCTTGGTATGGGGCAGCCGCA 2052
Qy 2041 GGACCTCAAAACCGGGCCACACCGCTTGTGGGGCTGATGCTTGGCTTCTCTGGGT 2100
Db 2053 GATGCTCGAGCCGGGGCCACAC---AGTGGGGCGGGAGTGGCTGGCTTCTTCTGGGG 2109
Qy 2101 GTTCTTGAGATCCCTCACTCTCTCTGATGGTGCCTGATGAGCAGCGTGCAGCAG 2160
Db 2110 ATTCTCGAGCATCCCTGACTCTCTATCTGATTTGGTGGCGCTCAGCAGCGACGCGAC 2169
Qy 2161 AGGAGCTTCTAGCTAGACACAAGTGGCTTACATCTGGGGCTGCCACCTTCTGGGACC 2220
Db 2170 AGGAATCTTGGCTAGACACAAGTGGCTGGGCTGGACCTGGGGCTGCCACCTTCTGGGACC 2229
Qy 2221 ACAAGCTATAGTCAGGACCTCCCTCTCTTCGCTGAAAGTGAACGCTGCCCTTGCC 2280
Db 2230 ACAAGCTACAGCAAGACCTCCCTCCCTCTCTCTGAAAGTGAAGCGGTTGCCCTGACC 2289
Qy 2281 CTGGGTAAAGGGGCGAGTGGTTTGGTGGCTTCCCTCCACCTTCTGCTGGATTTGCTG 2340
Db 2290 CTGGCCAAGAGGGGCGAGTGGTTTGGTGGATTTCTCAGCACCTTCTGATCTCTGCTG 2349
Qy 2341 CCAAGCCAGCCACATCCGCTCACCTGAGGGGCTCTAGCCAGCTGTGATGAGACCTCC 2400
Db 2350 CCAAGCCAGCCACATTCGCTAACCTGGGCTCCTCTAGCCACATGTGATGAACATCC 2409
Qy 2401 ATCTAAAGCCGGGAAATGACTGCCAGCCATGACGAGTCTCTCTGGAAC 2448
Db 2410 ATCTAGAGCTGGGCAATGACCACACTAGTGATATAGTGATCACTGGAAC 2457

RESULT 8
AK075384
LOCUS
DEFINITION Homo sapiens cDNA PSEC0074 fis, clone NT2RP2003050, highly similar
to Mus musculus mRNA for semaphorin W.
ACCESSION AK075384
VERSION AK075384.1 GI:22761435
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Ota,T., Nishikawa,T., Suzuki,Y., Kawai-Hio,Y., Hayashi,K.,
Ishii,S., Saito,K., Yamamoto,J., Wakamatsu,A., Nagai,T.,
Nakamura,Y., Nagahari,K., Sugano,S. and Isogai,T.
HRI human cDNA sequencing project
Unpublished

REFERENCE AUTHORS TITLE JOURNAL
2 (bases 1 to 2971)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (20-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
HRI human cDNA sequencing project; cDNA 5'- & 3'-end one pass
sequencing, clone selection and full insert sequencing; Helix
Research Institute (supported by Japan Key Technology Center etc.);
cDNA library construction; Institute of Medical Science, University
of Tokyo, Laboratory of Genome Structure, Human Genome Center.
FEATURES
source
1. 2971
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP2003050"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP2"
/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal
precursor cells after 2-weeks retinoic acid (RA)
induction."
103..2415
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC11584.1"
/db_xref="GI:22761436"
/translation="MPASARPRPGCQPTASPPFLLLAVLSPVSGRVRPSRPRD
LPISAEADSLTRFAVPHYNTSYVLLVPASHTLYVGARDTIFALSPFSGRRPRIDW
MVPEAHRQNRCKKQKEDCHNFVILAIANASHLLTCGTFADPKGCGVIDVSRFOQV
ERLESGRKCPFEPQAQSAVAGGVLYAATVKNYLTGTEPTITRAVRAAGDWRITDL
PSWLNAPFAVAAVSPAEGDEGDEIYFFFTETSRADFSEYRIKVRPRVACAG
LGGRTLQQRWTTFLKDLICPGPEHGRASSVLQDVALRPELGAGTPIFYGIFSSOW
EGATISAVCAFPOQDIRTVLNGPPELKHDCNRGLPVVDNDVPQPRGECITNNKLR
DVLGSLSDPLRVLTIRHPLMDRPFVADGHPLLVTDITATLRVAVHRTVSLGKEY
DFVLTGTEGHLHRAVRAQLSVLEDLALFPEPQPVENKLYHSLVLRSTETVQV
NTTNCGLQSCSECLAQDPVCAWSFRLDECAVHAGHRLVQDIESADVSLCPKEP
GERVFEVFPVATAHVLPVCPSSASVWHPQSGVTALTTPRRDGLVVVTPGAM
AYACEGGAHVAHVAVSLVWGSORDAPSRHTVAGLAGFFLGLAASLTLILGR
ROORRORRELLARDKVLGLDAGPSTTSYSDPPSPSPDERLPLALAKRSGFGGF
SPPELLDPCFSAHRLTGLAPLATCDETSI"
BASE COUNT 582 a 853 c 858 g 678 t
ORIGIN
Query Match 46.8%; Score 1875.2; DB 9; Length 2971;
Best Local Similarity 86.1%; Pred. No. 0;
Matches 2107; Conservative 0; Mismatches 323; Indels 18; Gaps 2;
Qy 1 GCCGAGGCCCGCGGAGTAGCGGTACTAAGTAGAGGCTGCTGGAGCGCCCGCCACCGCGCAC 60
Db 28 GCCGAGGCCAGTAGCCCGGGGCCCTTGAGCAGAGCGCGTAGCTTGGCGCGCACCGCGCGC 87
Qy 61 CAGCGGAGCCAGAGATGCTTGCAGGGCCCGAGGGCCCGCCCGCGGGCCCGCGCGCGCT 120
Db 88 CAGCGGAGCCAAAGATGCGCGGCTCTGCTCGCGGGCCCGCGGGTCCCGGGGAGCGCT 147
Qy 121 CCGCTCTTCCCTTCCCGCCCGCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 148 ACACCTCGCCCTTCCC-----GCTACTGCTGCTGCGGGGCTGCTGAGCGGC 192
Qy 181 CCGGTGTGCGGCGCGCTCCCGCGCTCAGTGCCCAAGAACCTCGCTGCCATCTCCGAGGCT 240
Db 193 CCGGTATCCGCGCGCTCCCGCGCTCGGTGCGCCAGAACCTCGCTGCCATCTCAGGCT 252
Qy 241 GACTCTATCTACCCCGGTTTGGAGGCTCTCATAGTACAAATTACTTGTCTCTCTTGTG 300
Db 253 GACTCTGTCTCACCCGGTTTCGAGTCCCTCACACATACAAATTACTTGTCTCTCTTGTG 312
Qy 301 GATCTGCTCCACACACTTTAGTGTGCGTGCAGGGATAGCATCTTGGCTTTTAAACCTC 360
Db 313 GATCTGCTCCACACACTTTATGTGTGGCGCCCGGACACCATCTTCTGCTTTATCCCTG 372

QY	361	CCCTTCTCTGGGGAAGACCCCGAAGATCGACTGGATGGTACCTGAGACTCACAGACAG	420
Db	373	CCCTTTCTCAGGGAGAGACCCCGCAGGATTTGACTGGATGGTTCTCCTGAGGCTCACAGACAG	432
QY	421	AACCTGAGGAAGAAAGGCAAGAGAGGAGCAATGTACAAATTTTATCAGATTTCTCGCC	480
Db	433	AACCTGTAGNAGAAAGCAAGAAAGAGGAGCAATGTACAAATTTTGTCCAGATTTCTCGCC	492
QY	481	ATTGTCAATGCCCTCTCACCTCCTCAGCTGCGGCACCTTCGCTTTTGTATCCGAAGTCCGGG	540
Db	493	ATTGCCAATGCCCTCTCACCTCCTCAGCTTTGTGGCACCTTCGCTTTTGTATCCGAAGTCCGGG	552
QY	541	GTTATTGATGTGTCCAGTTTCCAGCAGGTTGAAAGACTTGAGAGCGCGCGGGGAAATGT	600
Db	553	GTTATTGATGTGTCCAGTTTCCAGCAGGTTGAAAGACTTGAGAGTGGCGGGGGGAAATGT	612
QY	601	CTTTTGAAGCAGCTCAACGGTCAGCAGCTGTAATGGCTGGGGGCGCTCTCTACACCGCC	660
Db	613	CTTTTGAAGCAGCTCAGCGGTACGAGCTGTAAATGGCTGGGGGGGTCTCTATGCTGCC	672
QY	661	ACTGTGAAGAACHTTCTGGGACTGAGCCCATCATCTCCCGAGCTGTGGGTCCAGCTGAG	720
Db	673	ACTGTCAAAAACCTACCTGGGAGCGGAGCAATATATCACAGAGCAGTGGGTCTGTCGGAG	732
QY	721	GACTGATTCGACAGAGACCTTGTCTATCCTGGCTTAATGCTCCAGCCTTTGTGCGAGCT	780
Db	733	GACTGATTCGAGACATACCTTGGCTTCTCTGGCTGAAGCCCGCAGCCTTTGTGCGAGCC	792
QY	781	ATGGTCTCAGCCACGCTGAGTGGGGGATGAAGATGGAGAGCATGAAATCTTTTTC	840
Db	793	GTGGCCTTGAGCCACCGCAANTGGGGGATGAAGATGGAGAGCACGAAATCTACTTCTTC	852
QY	841	TTACGGAGACCTCCCGAGCTGTGGACTCCTATGAGCGCATCAAGTCCCAAGAGTGGCC	900
Db	853	TTTACGGAGACTTCCCGAGCAATTTGACTCATACGAGCGCATTAAGTCCCGAGGTGGCC	912
QY	901	CGAGTGTGTGGGGGACCTTGGGGCAGGAAGACCCCTTACAGAGATGGAGAGCTTT	960
Db	913	CGTGTGTGTGGGGGACCTTCGGGGCGGAGAGACCCCTCCAGCAGAGATGGACGAGTTT	972
QY	961	CTGAAGCTGACTGTGTGCCAGGGCCGAGCATGGCCGCTCCGGGCTTCGCGAG	1020
Db	973	TTGAAAGCTGACTGTGTCTGTCAGGGCCTGAGCATGGCGGGCTTCCAGTGTCTTCGAG	1032
QY	1021	GCTATGGCAGAGCTTGGCCTCAGCCTTGAGCGGGAACCCCATCTTTTATGGGATCTTT	1080
Db	1033	GATGTGTGTGTCTCGACCTGAGCTTGGGGCAGGAGCTCCCATCTTTTATGSCATCTTT	1092
QY	1081	TCTCTCCAGTGGAGAGAGCTGCCATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1140
Db	1093	TCTTCCAGTGGAGGGGGCTACTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1152
QY	1141	CGGGCAGTCTGAATGGTCCCTTTAGAGAGCTAAACATGACTGCAACAGGGAGCTGCCT	1200
Db	1153	CGGACAGTCTGAATGGTCCCTTACAGAGACTTAAACATGACTGCAACAGAGAGCTGCCT	1212
QY	1201	GTCATGGCAACAGAGTGGCCCGACCCAGACCTGGAGAGTGCATCGCCAAACATGAAG	1260
Db	1213	GTCGTGGACAATGATGTGCCCGACCCAGACCTGGAGAGTGCATCACCAACAACATGAAG	1272
QY	1261	CTCCAGCAGTTTGGATCTCATCTCTCCCTGCGCAGACCGGGTGTCTACCTTTATCAGAGAC	1320
Db	1273	CTCCGGCAGCTTTGGGTCTACTCTCTCTCCCTGAGCGGGTACTCACCTTTCATCCGGGAC	1332
QY	1321	CACCTCTCATGACAGCGCCCTGTCTCCGGCTGACGGCGCCCTCTGCTGTCTACTACA	1380
Db	1333	CACCCACTCATGACAGCGCCAGTGTTCAGCTGATGGCCACCCCTGCTGTCTACTACA	1392
QY	1381	GATACAGCCTATCTCAGAGTCTGTGGCCACAGGGGTGACCGCCTCTCAGGGGAAAGATAT	1440
Db	1393	GATACAGCCTATCTCAGAGTCTGTGGCCACAGGGGTGACCGCCTCTCAGGGGAAAGATAT	1452
QY	1441	GACGTGCTTACCTGGGACAGAGGATGGACACCTCCACCGGGCTGTGGCGATTTGGAGCT	1500

..

Db	1453	GATGTCTCTACCTGGGGACAGAGGATGGACACCTCCACAGCAGTGGCGATCGGAGCT	1512
QY	1501	CAGCTCAGTGTCTTTGGAGGATCTGGCTTGTTCCTCCAGAACACACAGCCGGTTGAGACATG	1560
Db	1513	CAGCTCAGCGTCTTTGAAGATCTGGCTTATTTCCAGAGCCACAGCCAGCTTGAACAATG	1572
QY	1561	AAATTGTACACCATTTGGCTTCTCGTGGGTCCCATACTAGGTGACACAAAGTGAACACC	1620
Db	1573	AAATTGTACACCATTTGGCTTCTCGTGGGTCCCATACTAGGTGACACAAAGTGAATACA	1632
QY	1621	AGCAACTGTGGCGCTTCCAGAGCTGCTCGGAGTGTATCTCTGCCCAGGACCCCGTGTGC	1680
Db	1633	ACCAACTGTGGCGCTTCCAGAGCTGCTCAGAGTGTATCTCTGCCCAGGACCCAGTCTGT	1692
QY	1681	GCTGTGAGCTTCCGGCTTGTGTGTGGCCACCGCGGAGCAGCAGCGGGATGGTT	1740
Db	1693	GCTGTGAGCTTCCGGCTTGTGTGTGGCCCATATGCGGGGAGCAGCGAGGTGGTC	1752
QY	1741	CAAGATATAGAGTACAGCGGATGTCTCTTTTGTGTCCAAAAGAACCTGGAGAACATCCC	1800
Db	1753	CAAGACATAGAGTACAGAGATGTCTCTTCTTGTGTCTTAAGAGAGCTTGAGAACGTCCA	1812
QY	1801	GTAGTGTTTGAAGTTCGGTGGCTACTGTGGGCCACAGTGGTCTCGCATGTTTCCCCAGT	1860
Db	1813	GTAGTGTTTGAAGTTCGGTGGCTACAGCTGCGCATGTGCTTGTGCATGTTTCTCCAAGC	1872
QY	1861	TCCTGCTGGCGCATCTGTGTGGCACACAGCCAGCCAGTGGAGTACTCGCTCACTCCCGG	1920
Db	1873	TCAGCATGGCATCTGTGTGTGGCACACAGCCAGTGGAGTACTCGCATCACTCCCGG	1932
QY	1921	AGGGATGAGTACTAGAGTGTGTGTGACCCAGGGGCCATGGGGCTTTATGCTTCGAGTGT	1980
Db	1933	CGGGATGAGTACTAGAGTGTGTGTGACCCAGGGGCCATGGGGCTTTATGCTTCGATGT	1992
QY	1981	CAGAGGGTGGAGCCCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2040
Db	1993	CAGAGGGTGGAGCCCGCATGT	2052
QY	2041	GGACCTCTAAACCGGGCCACACCGTGTGTGGGGCTGGAATGTTTGGCTTTCTCTCGGT	2100
Db	2053	GATGTCTCCAGCGCGGGGCCACAC - --AGTGGGGGGGAGTGGCTGGCTTCTTCTTGGGG	2109
QY	2101	GTCTTGGACATCCCTCACT	2160
Db	2110	ATTCTCGCAGCATCCCTGACT	2169
QY	2161	AGGGAGCTTCTAGCTAGACAAAGTGGCTTAGATCTGGGGCTCCACCTTCTTGGGACC	2220
Db	2170	AGGGAACTTCTGGCTAGAGACAAAGTGGGCTTGGACCTGGGGGCTCCACCTTCTTGGGACC	2229
QY	2221	ACAAGCTATAGTCAAGACCT	2280
Db	2230	ACAAGCTACGCAAGACCT	2289
QY	2281	CTGGTAAAGCGGGCAGTGGTTTGTGGCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2340
Db	2290	CTGGCAAGAGGGGAGTGGCTTTGTGGATTTCTCAACACCTTCTCTCTCTCTCTCTCTCT	2349
QY	2341	CCAAGCCAGCCACATCCGGCTCACTGGGGGCGCTCTTACGACAGTGTGTATGAGACCTCC	2400
Db	2350	CCAAGCCAGCCACATTCGGCTAACTGGGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2409
QY	2401	ATCTAAAGCGGGGAAATGACTGCCAGCATAGTGTATAGTGTACTCTGGAAC	2448
Db	2410	ATCTAGAGCTGGGCAATGACCCACTAGTGTATAGTGTACTCTGGAAC	2457

RESULT 9
BC038411
LOCUS

2948 bp mRNA linear PRI 02-OCT-2002
Homo sapiens, sema domain, immunoglobulin domain (Ig),
transmembrane domain (TM), and short cytoplasmic domain.

(semaphorin) 4F, clone MGC:34479 IMAGE:5179587, mRNA, complete cds.

ACCESSION

BC038411

VERSION

1 GI:23468353

KEYWORDS

MGC.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS

Strausberg, R.

TITLE

Direct Submission

JOURNAL

Submitted (01-OCT-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK

NIH-MGC project URL: <http://mgc.nci.nih.gov>

COMMENT

Contact: MGC help desk

Email: cgaps-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,

Yoon, V.S., Kowls, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,

Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/ILLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 51 Row: p Column: 3

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 19923278.

FEATURES

source

1..2948

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:34479 IMAGE:5179587"

/tissue_type="Brain, Lung, Testis, adult, pooled whole"

/clone_lib="NIH_MGC_115"

/lab_host="DH10B"

/note="vector: pCMV-SPORT6"

51..2363

/codon_start=1

/product="sema domain, immunoglobulin domain (Ig),

transmembrane domain (TM) and short cytoplasmic domain,

(semaphorin) 4F"

/protein_id="AAH38411.1"

/db_xref="GI:23468354"

/db_xref="LocusID:10505"

/translation="MPASARPRPGQPTASFPPLLLAVLSGPVSRVPRSPRTS

LPISAECDKTRFVPHYNYSVLLVDPASHTLVGARDTIFALSPFSGRPRRDIW

MVPEARQNCRRKKEDECNFVQILAIANASHLLTCGTFAPDKGVLDVSRFQV

ERLESRGKCPFPAPORSAAMAGVLYAATVKNLYLTPITRANGRAEDWRTDL

PSWLNAPAFVAALSPAEGWDEGDDIYVFETESRAFDISVRIKPVVRVACGD

LGGRKTQQRWTFKLADLLCPGHEGRASSVLQDVAVLPDELGACTPIFYGLFSOW

EGATISAVCAFRQDITVYNGPRELKHDCNGLPVDNDVQPRPGEITNNMKLR

HFGSSLSRVRVTFIDHPLMDRPFVPPADLLVTTDTAYLRVVAHVTSLSGKEY

DVLYGTDEGHLHRAVRIGLAQLSVLEDALEFPFVNNMKLYHSLVLSRTEVTV

NTYNGRLQSCSICILAQDPVCAWSFRLECAVAGEHRLQVODIESADYSSLCRKP

GERPVPEVPAVAHVLPSPSSAWSCVHQPGVATLTPRROGLVTVVTPGAMG

AVACEQEGGAHVVAAYSLVWGSRQDAPSRATVAGLAGFLGILAAASLTLLIGR

RQRRRLRELLAKVDGLDGLAPSGTTSYQDPPSPSPEDERLPLALARKSGFGFG

SPFFLLDPCSPAHIRLTGAPLATCDTSI"

602 a 834 c 836 g 676 t

BASE COUNT

ORIGIN

Query Match 46.6%; Score 1867.8; DB 9; Length 2948;
Best Local Similarity 86.4%; Pred. No. 0;
Mismatch 0; Mismatches 312; Indels 18; Gaps 2;

26 TAAGTAGAGGCTGCTGACGCGCCGCCACCGGCACAGGGGAGCCAGAGATGCTTGCCA 85
1 TGACAGAGGCGCTAGCTTGGCGCGACCCGCGGAGGAGCAAGATGCGGCGCT 60
86 GGGCCGAGGCGCCCGCGGCGGCGGCGGCTCCGGTCTTTCCCTTCCCGCGCGC 145
61 CTGTCGCGCGCCCGCGGCTCCCGGAGCCTACAGCCTCCCGCTTCC----- 112
146 TGTGCTGCTGCTGCTGCTGCGGATCTAAGCGCCCGCGGCTGCGGCGCGCTCCCGCT 205
113 -----GCTACTGCTGCTGCGGCTGCTGAGCGCGCGGATCCGCGCGGCTCCCGCT 165
206 CAGTGCACAGAACCTCGCTGCCATCTCCGAGGCTGACTCTCTATCTCACCGGTTTGAC 265
166 CGGTGCCAGAACCTCGCTTCCAATCTCTGAGGCTGACTCTGCTCACCGGTTGCGAG 225
266 CGTCTCATAGTCAATTAATCTGCTCTCTCTGATGCTGCTGCTGCTGCTGCTGCTGCT 325
226 TCCCTCACACATAACAATTAATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 285
326 TCGGTGACGCGGATAGCATCTTGGCTTTTAAACCTTCCCTTCTCTGCGGAAAGACCCGAA 385
286 TTGGCGCCCGGACACCATCTTGGCTTTTATCCCTTCCCTTCTCTGAGGGAGAGCCCGCA 345
386 GGATCGACTGGATGCTACCTGAGACTCACAGACAGAACTGCAGGAAGAAAGGCAAGAAAG 445
346 GGATGCTGATGGTGTCTGAGGCTCACAGACAGAACTGTAGGAAGAAAGGCAAGAAAG 405
446 AGGACGAATGTCAATTTTATCCAGATTCGCGCAATTCGCAATTCCTCTCACCTCTCA 505
406 AGGACGAATGTCAATTTTATCCAGATTCGCGCAATTCGCAATTCCTCTCACCTCTCA 465
506 CGTGGCGGACCTTGGCTTTTGAATCCGAGGCGGGGTTATGATGCTGCTGCTGCTGCTGCT 565
466 CTTGTGCGACCTTGGCTTTTGAATCCGAGGCGGGGTTATGATGCTGCTGCTGCTGCTGCT 525
566 AGTTTGAAGACTTTGAGAGGCGCGGGGAAATGCTCTTTTGGCCAGCTCAACGGGTCAG 625
526 AGTTTGAAGACTTTGAGAGGCGCGGGGAAATGCTCTTTTGGCCAGCTCAACGGGTCAG 585
626 CAGCTGTAATGGCTGGGGGCTCTCTACACCGGCACTGTGAAGAACTTCTGCTGGGACTG 685
586 CAGCTGTAATGGCTGGGGGCTCTCTATGCTGCACTGTGAAGAACTTCTGCTGGGACTG 645
686 AGCCCATCATCTCCCGAGCTGTGGTGTGAGCTGAGGACTGGATTTCGAACAGAGACCTGT 745
646 AGCCCATCATCTCCCGAGCTGTGGTGTGAGCTGAGGACTGGATTTCGAACAGAGACCTGT 705
746 CATCTCTGGCTTAATGCTCCAGCTTTTGTGCGAGCTATGGTCTTACAGGAGACCTCCGAGTGT 805
706 CTTCTCTGGCTTAATGCTCCAGCTTTTGTGCGAGCTATGGTCTTACAGGAGACCTCCGAGTGT 765
806 GGGATGAAGTGAAGAGAGTGAATCTTTTTCCTACAGGAGACCTCCGAGTGTGTTGG 865
766 GGGATGAAGTGAAGAGAGTGAATCTTTTTCCTACAGGAGACCTCCGAGTGTGTTGG 825
866 ACTCTTATGAGCGATCAAGTCTCCCAAGAGTGGCGGCGGCTGTGTGCGGGGACCTTGGGG 925
826 ACTCATAGGAGCGATTAAGTCTCCAGCGGCTGGCGGCTGTGTGCGGGGACCTTGGGG 885
926 GCAGGAAGACCTTTTACAGAGAGTGGAGCGAGCTTTTCTGAAGCTGAGCTGCTGTGCCAG 985
886 GCCGAAGACCTTTTACAGAGAGTGGAGCGAGCTTTTCTGAAGCTGAGCTGCTGTGCCAG 945
986 GGCCCGAGCATGGCGGCGCTCCCGGGTCTGCGAGGCTATGGCAGAGCTTCCGCGCTCAGC 1045
946 GGCTGAGCATGGCGGCGCTCCAGTGTCTGCGAGGATGTGCTGTGCTGCTGCTGCTGCTGCT 1005
1046 CTGGAGCGGAGACCTTCTTTTATGGATCTTTTCTCCAGTGGGAGGAGCTGCTGCTGCTGCT 1105
1006 TTGGGCGAGGAGCTCCCATCTTTTATGGATCTTTTCTCCAGTGGGAGGAGGCTGCTGCTGCT 1065

QY 1106 TCTCTGCTGTGTGCTTCCAGACCCCAAGACATCCGGGCGAGTGTGTAATGTCCTCTTA 1165
Db 1066 TCTCTGCTGTGTGCTTCCAGACCAAGACATTCGGAGAGTGTGTAATGTCCTCTTA 1125
QY 1166 GAGAGCTAAACATCAGTGCACAGGGGAGTGCCTGTCTATGACACAGAGGTGCCCCAGC 1225
Db 1126 GAGAATAAATCAGTGCACAGAGGAGTGCCTGTGCTGGACAATGATGTGCCCCAGC 1185
QY 1226 CCAGACCTGGAGTGCATCGCAACAACATGAAGCTCCAGAGTTTGGATCCTCACTCT 1285
Db 1186 CCAGACCTGGAGTGCATCACCACAACATGAAGCTCCGGCACTTTGGCTCATCTCT 1245
QY 1286 CCTCTGCAGACCGCTGTCTACCTTTATCAGAGACCAACCTCTCATAGGAGGCCGTGT 1345
Db 1246 CCTCTGCCTGACCGCTGTCTACCTTTATCAGAGACCAACCTCTCATAGGAGGCCGTGT 1305
QY 1346 TCCCGGCTGACGGCGCCCGCTGCTGCTACATACAGATACAGCCTATCTCAGAGTCGTGG 1405
Db 1306 TTCCAGCTGTATGGCCACCCCGCTGCTGCTACATACAGATACAGCCTATCTCAGAGTCGTGG 1365
QY 1406 CCACAGGGTGACCAAGCCTCTCAGGGAAGAATATGAGTGTCTACCTGGGGACAGAGG 1465
Db 1366 CCACAGGGTGACCAAGCCTCTCAGGGAAGAATATGAGTGTCTACCTGGGGACAGAGG 1425
QY 1466 ATGGACACTCCACCGGGGTGTGCGCATTTGGAGCTCAGTGTCTTTGGAGGATCTGG 1525
Db 1426 ATGGACACTCCACCGGGGTGTGCGCATTTGGAGCTCAGTGTCTTTGAAGATCTGG 1485
QY 1526 CTTCTGTTCCAGAACCAAGCGGTTGAGAGCATGAAATTTACCAAGATTTGGCTCTCTGG 1585
Db 1486 CTTCTGTTCCAGAACCAAGCGGTTGAGAGCATGAAATTTACCAAGATTTGGCTCTCTGG 1545
QY 1586 TGGGCTCCCATCTAGTGTGACACAAGTGAACACCAAGCAACTGTGCGCTCTCCAGAGCT 1645
Db 1546 TTGGCTCCCGTACTGAGTGTGACACAAGTGAATACAACTGTGCGCTCTCCAGAGCT 1605
QY 1646 GCTCGAGTGTATCTGCGCAGGACCCCGTGTGCGCTGAGCTTCCGGTGTGATGTT 1705
Db 1606 GCTCAGAGTGTATCTGCGCAGGACCCCGTGTGCGCTGAGCTTCCGGTGTGATGTT 1665
QY 1706 GTGTGGCCACCGCGGCGAGCACCGGATGTTCAAGATATAGAGTACAGGATGCT 1765
Db 1666 GTGTGGCCACCGCGGCGAGCACCGGATGTTCAAGATATAGAGTACAGGATGCT 1725
QY 1766 TTCTTTGTGTCAAAAGAACCTGGAGAACATCCCGTGTGTTGAAGTTCCCGTGGCTA 1825
Db 1726 CCTCTTTGTGCTTAAGAGCCTGGAGAACCTGTCAGTAGTGTGTTGAAGTTCCCGTGGCTA 1785
QY 1826 CTGTGGGCGACGTGTGCTGCGATGTTCCCGAGTGTGCTGCGGATCCTGTGTGTGGC 1885
Db 1786 CAGCTGCGCATGTGCTGTGCGATGTTCTCCAAGCTCAGCATGGGCATCCTGTGTGTGGC 1845
QY 1886 ACCAGCCAGTGGAGTGTGCTGCTCACTCCCGGAGGATGGACTAGAGTGGTGTGA 1945
Db 1846 ACCAGCCAGTGGAGTGTGCTGCTCACTCCCGGAGGATGGACTAGAGTGGTGTGA 1905
QY 1946 CCCAGGGCCATGGGGCTTATGCTTGCAGTGTTCAGGAGGTGGAGCCCGCGCTGG 2005
Db 1906 CCCAGGGCCATGGGGCTTATGCTTGCAGTGTTCAGGAGGTGGAGCCCGCGCTGG 1965
QY 2006 TGGCTGTATTAGCTTGTGTGGGCGACCGAGCGGGACCTCAAAACCGGGCCACACCG 2065
Db 1966 TAGCAGTTACAGCTTGTGTGGGCGACCGAGAGATGCTCCGAGCGGGGCCACAC - 2023
QY 2066 TTGTGGGGGTGGATTGTGCTGCTTTCCTGCTGGGTGTTCTTGCAGCATCCCTCACTCNC 2125
Db 2024 -AGTGGGGCGGGACTGCTGCTTCTTCTTGGGATTTCTCGCAGCATCCCTGACTCPCA 2082
QY 2126 TCCTGATTGTGCGCGTTCAGGAGCGTTCGGCGACAGAGGAGTCTTAGCTAGAGACAAGG 2185
Db 2083 TTCTGATTGTGCGCGTTCAGGAGCGTTCGGCGACAGAGGAGTCTTAGCTAGAGACAAGG 2142
QY 2186 TGGGCTTAGATCTGGGGGCTCCACCTTCTGGGACCAAGCTATAGTACGAGACCTCCT 2245

Db 2143 TGGGCTGAGACCTGGGGCTCCACCTTCTGGGACCAAGCTATACGCCAAGACCTCCCT 2202
QY 2246 CTCCTTCCCTGAAGATGAAGGCTGCCCTTGGCCCTGGGTAAAGCGGCGACTGTGTTTG 2305
Db 2203 CCCCTCTCTCTGAAGATGAGCGTTCGCCCTGGCCCTGGCCAAGGGGCACTGGCTTG 2262
QY 2306 GTGGCTTCCCTCCACCTTCTCTGCTGGATTTGCCCCAAGCCAGCCACATCCCGCTCA 2365
Db 2263 GTGGATTCACACACCTTCTCTGCTGATCTTCCCAAGCCAGCCACATTCGGCTAA 2322
QY 2366 CTGGGGCGCTTACGCCAGCTGTGATGAGACCTCATCTAAAGCCGGGGGAAATGACTGC 2425
Db 2323 CTGGGGCTCTCTGATGAAACATCATCTAGACTGGGCAATGACCCT 2382
QY 2426 CAGCCATGAGCAGTCTCTGGAC 2448
Db 2383 AGTGATAAGTACACTGGAAC 2405
RESULT 10
AX003081
LOCUS AX003081 2893 bp DNA linear PAT 24-AUG-2000
DEFINITION Sequence 1 from Patent EP0933425.
ACCESSION AX003081
VERSION AX003081.1 GI:9926964
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Michalovich,D. and Doe,T.R.
TITLE Semaphorin family polypeptides and polynucleotides
JOURNAL Patent: EP 0933425-A 1 04-AUG-1999;
SMITHKLINE BEECHAM PLC (GB)
LOCATION/Qualifiers
FEATURES
source
1..2893
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 579 a 825 c 842 g 647 t
ORIGIN
Query Match 41.8%; Score 1673.6; DB 6; Length 2893;
Best Local Similarity 82.2%; Pred. No. 0;
Matches 2012; Conservative 0; Mismatches 319; Indels 117; Gaps 3;
QY 1 GCCGAGGCCCGCGCAGTAGAGGCTACTAAGTAGAGGCTGCTGGACGCGCGCCACCGCGCAC 60
Db 35 GCCGAGGCCCGCGCAGTAGAGGCTGCTGGACGCGCGCTAGCTTGGCGCGCACCGCGCGC 94
QY 61 CAGCGGAGCCAGAGATGCTTGGCAGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCT 120
Db 95 CAGCGGAGCCAAAGATGCCGCGCTCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCT 154
QY 121 CCGGCTTCTTCCCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Db 155 ACAGCTCGCCCTTCCC-----GCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 199
QY 181 CCGGCTGCTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 200 CCGGCTGCTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 259
QY 241 GACTCTCTATCTACCCCGGTTTGCAGCGCTCTCATAGTACAATTAATCTGCTGCTCTCTG 300
Db 260 GACTCTCTGCTACCCCGGTTTGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 319
QY 301 GATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 320 GATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 379
QY 361 CCCCTTCTGCGGGAAGACCCCGAGGATGCTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

[illegible]

RESULT 11
BD177663
LOCUS
DEFINITION
ACCESSION
VERSION


```
QY 661 ACTGTGAAGAACTTCTCGGGAGCTGAGCCCATCATCTCCGAGCTGTGGGTGAGAGCTGAG 720
Db 581 ACTGTGAAGAACTTCTCGGGAGCTGAGCCCATCATCTCCGAGCTGTGGGTGAGAGCTGAG 640
QY 721 GACTGATTCGAGACAGAGACCTTGTATCTCGCTTAATGCTCCAGCCCTTTCTCGAGCT 780
Db 641 GACTGATTCGAGACAGATACCTTGTGCTTCTCGCTGAAGCCGCCAGCCCTTTCTCGAGCC 700
QY 781 ATGGCTCTGAGCCAGCTGAGTGGGGGATGAAGATGGAGACGATGAATCTTTTTC 840
Db 701 GTGGCTCTGAGCCAGCCGAATGGGGGATGAAGATGGAGACGATGAATCTTACTTCTC 760
QY 841 TTCAGGAGACCTTCGGAGTGTGGACTCCTATGAGCGCATCAAGGTCCTCCAGAGTGGCC 900
Db 761 TTTACGGAGACTTCCGAGCATTTGACTCATACGAGCGCATTTAAAGTCCACGGGTGGCC 820
QY 901 CGAGTGTGTGGGGGAGCTTCGGGCGAGAGACCCCTTCACAGAGATGGAGAGCTTT 960
Db 821 CGTGTGTGTGGGGGAGCTTCGGGCGCGGAAGACCCCTCCAGAGATGGAGAGCTTT 880
QY 961 CTGAAGGCTGACCTGCTGTGCCAGGGCCCGAGCATGGCCGGGCTTCGGGCTTCTGCAG 1020
Db 881 TTGAAGAGCTGACCTGCTGTGCCAGGGCTGAGCATGGCCGGGCTTCAGTGTCTCGCAG 940
QY 1021 GCTATGGCAGAGCTTCGGGCTCAGGCTGGAGCGGGAACCCCATCTTTTATGGGATCTTT 1080
Db 941 GATGTGTGTGCTTGTGAGCTTGGGGGAGGAGCTTCCCATCTTTTATGGCATCTTT 1000
QY 1081 TCTCCAGTGTGGAAGGAGCTGCCATCTCTGCTGTGTGCTGTGCTTCCGACCCCAAGACATC 1140
Db 1001 TCTTCCAGTGTGGAAGGAGCTACTATCTGCTGTGTGCTGTGCTTCCGACCAAGACATTT 1060
QY 1141 CGGGCAGTGTCTGAATGGTCCCTTTAGAGAGCTTAAACATGACTGCAACAGGGGACTTGCCT 1200
Db 1061 CGGACAGTGTCTGAATGGTCCCTTCAGAGAACTTAAACATGACTGCAACAGAGGACTGCCT 1120
QY 1201 GTCATGGCAACAGAGTGGCCAGCCAGCCAGACTGTGAGAGTGCATGCCAACAAATGAAG 1260
Db 1121 GTCGTGGACAATGATGTGCCCCAGCCAGACCTGGAGAGTGCATCACCACAAATGAAG 1180
QY 1261 CTCAGCAGTGTGGATCTCTACTCTCCCTGCGAGACCGGCTGCTACCTTTTATCAGAGAC 1320
Db 1181 CTCGGGCACTTTGGCTCATCTCTCTCCCTGCTGACCGGCTACTCACCTTCTATCGGGAC 1240
QY 1321 CACCTCTCATGGACAGCCGCTGTTCGGGTGACGCGCCGCTGCTGTGCTACTACTACA 1380
Db 1241 CACCCACTCATGGACAGCCAGTGTTCAGCTGATGGCCACCCCTGCTGCTACTACA 1300
QY 1381 GATACAGCTATCTCAGAGTCTGTGGCCACAGGGTGACAGGCTCTCAGGGAAAGATAT 1440
Db 1301 GATACAGCTATCTCAGAGTCTGTGGCCACAGGGTGACAGGCTCTCAGGGAAAGATAT 1360
QY 1441 GAGCTGCTTACTCTGGGACAGAGGATGACACCTCCACCGGCTGTGGGCAATTTGAGCT 1500
Db 1361 GATGTGCTTACTCTGGGACAGAGATGGACACCTCCACCGAGAGTGGGATTCGGAGCT 1420
QY 1501 CAGCTCAGTGTCTTGAGGATGTGGCTTGTTCAGAACCCAGCCGCTTGAGAGCATG 1560
Db 1421 CAGCTCAGGCTTTTGAAGATCTGGCTTATTTCCAGAGCCACAGCCAGTTCAGAACATG 1480
QY 1561 AAATGTACCAGATTTGGCTCTGTGGGCTCCCATAGTGAAGTGCACAGTGAACACC 1620
Db 1481 AAATGTACCAGTGGCTCTGTGGGCTCCCATAGTGAAGTGCACAGTGAATACA 1540
QY 1621 AGCAACTGTGGCGGCTCCAGAGTGTCTGGAGTGTATCTGGCCAGGACCCCGTGTGC 1680
Db 1541 ACCAAGTGTGGCGGCTCTCAGAGTGTCTCAGAGTGTATCTGTGGCCAGGACCCAGTCTGT 1600
QY 1681 GCCTGAGCTTCCGGCTTGATCTTGTGTGGCCCAAGCGGCGGAGCACCGCGGATGGTT 1740
Db 1601 GCCTGAGCTTCCGGCTTGATCTTGTGTGGCCCAAGCGGCGGAGCACCGGAGGTTGGTC 1660
QY 1741 CAAGATATAGATCAGCGGATCTCTCTTTGTGTCCAAAGAACAGCTGGAGAACATCC 1800
```

```
Db 1661 CAAGACATAGAGTCAGCAGATGTCTCTCTTTGTGTCTTAAAGAGCCTGGAGAACGTCCA 1720
QY 1801 GTAGTGTTTGAAGTTCCCGTGGCTACTGTGGGCCAGCTGTGGCATGTTTCCCCACT 1860
Db 1721 GTAGTGTTTGAAGTTCCCGTGGCTACAGCTGGCATGTGGTTCGTCATGTTCTTCCAAGC 1780
QY 1861 TCTGCTGGGCACTCTGTGTGGCACCAGCCAGCTGGAGTGACTGCGCTACTCCTCCCGG 1920
Db 1781 TCAGCATGGGCATCTGTGTGGCACCAGCCAGCTGGAGTGACTGCACTACCCCCCGG 1840
QY 1921 AGGATGACTAGAGGTGTGTGATCCCGAGGGCCATGGGGCTTATCTTTCGAGTGT 1980
Db 1841 CGGGATGACTGGAGGTGTGTGATCCCGAGGGCCATGGGGCTTATGCTTGAATGT 1900
QY 1981 CAGGAGGTGGAGCGCCGCTGTGGCTGTATAGCTTGTGTGGGCGAGCAGCGG 2040
Db 1901 CAGGAGGTGGGCGAGCCCATGTGTAGCAGCTTACAGTTGGTATGGGGCAGCAGCA 1960
QY 2041 GGACCTCAAAACCGGGCCACACCGTTGTGGGGCTGGATTGGTTTCTCTGGGT 2100
Db 1961 GATGCTCCGAGCGGGCCACAC---AGTGGGGCGGAGCTGGCTGGCTTCTTGGGG 2017
QY 2101 GTTCTTGAGCATCCCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2160
Db 2018 ATTCTCGCAGCATCCCTGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2077
QY 2161 AGGAGCTTCTAGCTAGAGACAAGTGGGCTTGTAGATCTGGGGCTCCACCTTCTGGGACC 2220
Db 2078 AGGAACTTCTGGCTAGAGACAAGTGGGCTTGGACCTGGGGCTCCACCTTCTGGGACC 2137
QY 2221 ACAAGCTATAGTCAGGACCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2280
Db 2138 ACAAGCTACAGCAAGACCCCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2197
QY 2281 CTGGGTAGCGGGGAGTGGTTTGGTGGCTTCCCTCCACCTTCTCTCTCTCTCTCTCTCTCTCTCT 2340
Db 2198 CTGGCAAGAGGGGAGTGGCTTGGTGGATTTCTCACCACCTTCTCTCTCTCTCTCTCTCTCTCTCT 2257
QY 2341 CCAAGCCAGCCACATCCGGCTACTGGGGCGCTCTAGCCACGCTGTGATGAGACCTCC 2400
Db 2258 CCAAGCCAGCCACATTCGGCTACTGGGGCTCTCTAGCCACATGTGATGAACATCC 2317
QY 2401 ATCTAAAGCCGGGAAATGACTGCGAGCCATGAGCAGTCTCTGGAAC 2448
Db 2318 ATCTAGAGCTGGGCAATGACCACCTAGTGTATAGTGTATCACTCTGGAAC 2365
```

RESULT 13

```
AX003083
LOCUS 2894 bp DNA linear PAT 24-AUG-2000
DEFINITION Sequence 3 from Patent EP0933425.
ACCESSION AX003083
VERSION AX003083.1 GI:9926965
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Michalovich, D. and Doe, T.R.
TITLE Semaphorin family polypeptides and polynucleotides
JOURNAL Patent: EP 0933425-A 3 04-AUG-1999;
SMITHKLINE BEECHAM PLC (GB)
FEATURES
source 1..2894
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 580 a 824 c 841 g 646 t 3 others
ORIGIN
```

Query Match 41.8%; Score 1673.6; DB 6; Length 2894;

Qy	2161	AGGAGCTTCTACGTAGACAAAGTGGGCTTAGATCTGGGGCTCCACCTTTCTGGGACC	2220
Db	2078	AGGAACTTCTGCTAGAGACAAAGTGGGCTTGACCTGGGGCTCCACCTTCTGGGACC	2137
Qy	2221	ACAAGCTATAGTCAGAGACCTCCCTCTCTCTGCCCTGAAGATGAAGCGTGCCTCTGGGC	2280
Db	2138	ACAAGCTACAGCCAAAGACCTCCCTCCCTCTCTCTGAAGATGAGCGGTGGCGCTGGGC	2197
Qy	2281	CTGGTAAAGGGGGCAGTGGTTTTGGTGGCTTCCCTCCACCTCTCTGCTGGATCTCTGC	2340
Db	2198	CTGCCAAGAGGGGCAGTGGCTTTGGTGGATCTCACCACCTTCTCTGCTTGATCTTGC	2257
Qy	2341	CAAAGCCAGCCACATCCGGCTCACTGGGGCGCTCTAGCCACAGTGTGATGAGACCTCC	2400
Db	2258	CAAAGCCAGCCACATCCGGCTAACTGGGCTCTCTAGCCACATGTGATGAACATCC	2317
Qy	2401	ATCTAAGCGGGGAAATGACTGCCAGCATGAGCAGTCTCTGGAAC	2448
Db	2318	ATCTAGAGCTGGCAATGACCACTAGTGTATTAAGTGATCACTGGAAC	2365
RESULT 14			
LOCUS	BD177664	2894 bp	DNA linear PAT 16-APR-2003
DEFINITION	SBSEMN1 polypeptides and polynucleotides.		
ACCESSION	BD177664		
VERSION	BD177664.1 GI:30014926		
KEYWORDS	JP 2002281988-A/2.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1. (bases 1 to 2894)		
TITLE	Michalovich,D., Doe,T.R. and Hayes,P.D.		
JOURNAL	SBSEMN1 polypeptides and polynucleotides		
COMMENT	Patent: JP 2002281988-A 2 02-OCT-2002; SMITHKLINE BEECHAM PLC OS Homo sapiens (human) PN JP 2002281988-A/2 PD 02-OCT-2002 PF 20-DEC-2001 JP 2001388219 PR 28-JUL-1998 GB 9816423:9,30-JAN-1998 GB 98300693:3 PR 30-JUL-1998 GB 9816676:2 PI DAVID MICHALOVICH, TRUDY R DOE, PHILIP D HAYES PC C12N15/09,A61K45/00,A61K48/00,A61P21/04,A61P25/18,A61P29/00, PC A61P31/12,A61P35/00,A61P37/00,C07K14/47,C07K16/18,C12N1/15,PC C12N1/19, PC C12N1/21,C12N5/10,C12P21/02,C12Q1/02,C12Q1/68,G01N33/15,G01N33/PC 50, PC G01N33/53,C12N15/00,C12N5/00 CC SBSEMN1 polypeptides and polynucleotides FH Key Location/Qualifiers FT source 1. 2894 FT Location/Qualifiers /organism='Homo sapiens (human)'. /mol_type='genomic DNA' /db_xref='taxon:9606'		
BASE COUNT	580 a	824 c	841 g 646 t 3 others
ORIGIN			
Query Match	41.8%; Score 1673.6; DB 6; Length 2894;		
Best Local Similarity	82.2%; Pred. No. 0;		
Matches 2012:	Conservative	0; Mismatches	319; Indels 117; Gaps 3;
Qy	1	GCGAGCGCCGCGAGTAGCGGTACTAGTAGAGGCTGTGAGACGCGCCACCCGCGAC	60
Db	35	GCGAGCGCCAGTAGCCCCGGGCGCTGAGCAGAGCGCGTAGCTTGGCGCGCACCCGCGC	94

Db	2138	ACAAGCTACAGCAAGAACCTCCCTCCCTCCCTCTCCTGAAGATGAGCGGTTCGCGCTGGCC	2199
Qy	2281	CTGGTAAAGCGGCGAGTGGTTTTGGTGGCTTCCTCCACCCCTCCTGCTGGGATTTCTTGC	2340
Db	2198	CTGCCAAGAGGCGAGTGGCTTTGGTGATTTCTACCCACCCCTCCTGCTTGGCTTGCCTTGC	2257
Qy	2341	CTAAGCCAGCCACATPCCGGCTACATGCGGGCGCTCTAGCCACGTTGATGATGAGACCTC	2400
Db	2258	CAAAGCCAGCCACATTCGGCTAACTGGGGCTCTCTAGCCACATGTGATGAACATCC	2317
Qy	2401	ATCTAAGCCGGGAAATGACATGCCAGCCATGAGCAGTCTCTGGAC	2448
Db	2318	ATCTAGAGCTGGGCAATGACCACTAGTGATTAAGTGATCACTGGAAC	2365
RESULT	15		
E35444			
LOCUS	E35444	2894 bp	DNA
DEFINITION	SBSEMN1 polypeptide and polynucleotide.	linear	PAT 18-JUN-2001
ACCESSION	E35444		
VERSION	E35444.1	GI:13019082	
KEYWORDS	JP 199225778-A/2		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 (bases 1 to 2894)		
TITLE	David, M., Tudy, R.D. and Philip, D.H.		
JOURNAL	SBSEMN1 polypeptide and polynucleotide		
COMMENT	Patent: JP 199225778-A 2 24-AUG-1999; SMITHKLINE BEECHAM CORP PUBLIC LTD CO OS Homo sapiens (human) PN JP 199225778-A/2 PD 24-AUG-1999 PF 27-OCT-1998 JP 1998306049 PR 28-JUL-1998 GB 9816423; 9, 30-JAN-1998 GB 98300693; 3, PR 30-JUL-1998 GB 9816676; 2 PI DAVID MIKAROBITU, TUDY R DOE, PHILIP D HAYES PC C12N15/09, A61K45/00, C07K14/47, C12P21/02, C12D1/68, G01N33/15, PC G01N33/53 PC C12N15/00 CC FH Key FT Location/Qualifiers FT Location/Qualifiers 1. .2894 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"		
FEATURES	source		
BASE COUNT	580 a 824 c 841 g 646 t	3 others	
ORIGIN			
Query Match	41.8%	Score 1673.6;	DB 6; Length 2894;
Best Local Similarity	82.2%	Pred. No. 0;	
Matches 2012;	Conservative	0; Mismatches 319;	Indels 117; Gaps 3;
Qy	1	GCGAGGCCCGCGAGTAGGGGTACTAAGTAGAGGCTGCTGGACGGCCGCCACCCCGCGCAC	60
Db	35	GCGAGGCCCACTAGCCCCGGGGCCCTGAGCAGAGCGCGTAGCTTGGCGCGCACCCCGCGGC	94
Qy	61	CAGCGGAGCGCAGAGATGCTTTCAGGGCCGAGCGCCGCCGCCGCCGCCGCCGCCGCCCT	120
Db	95	CAGCGGAGCGCAAGATGCGCGGCTCTGCTGCGCGGCCGCCGCCGCCGCCGCCGCCCT	154
Qy	121	CCGGTCTTTCCCTTCCCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	180
Db	155	ACAGCCTCGCCCTTCCC-----GCTACTGCTGCTGGCGGTGCTGAGCGGC	199
Qy	181	CCGGTGTGCGGCGCGTCCCGCTCAGTGCCAGAACCTCTGCTGCGCCATCTCCGAGGCT	240
Db	200	CCGGTATCCGCGCGCGTCCCGCTCGGTGCCCAAGACCTTCGCTTCCAACTCTGAGGCT	259

Qy	241	GACTCCTPANTCAACCCGGTTTGGCAGCGTCTCATACGTACAAATTA	CTCTGCTCTCCTTTGTG	300
Db	260	GACTCCTGTCTACACCCGGTTTGGCAGTCCCTCACACATACAATTA	CTCTGCTCTCCTTTGTG	319
Qy	301	GATCCTGCCCTCCACACACTTTACGTGGTGCACGGGATAGCATTC	TTGCGCTTTAAACCCCTC	360
Db	320	GATCCTGCCCTCCACACACTTTATGTTGGCGCCGGACACCAATTC	TTGCGCTTTATCCGCTG	379
Qy	361	CCCTTCTCTGGGGAAGACCCCGAAGGATCGACTGGATGGTACCTG	TGAGACTTCACAGACAG	420
Db	380	CCCTTCTCAGGGGAGACCCCGCAGAGTTGACTTGGATGGTTCTG	TGAGGCTTCACAGACAG	439
Qy	421	AACTGCAAGGAAGAAAGCAAGAAAGAGACGAATGTCAACAATTT	TATACCAAGATTCTCGCC	480
Db	440	AACTGTAGGAAGAAAGCAAGAAAGAG	-----	466
Qy	481	ATTGTCAATGCCTCTCAGCTCCTCACGTGGCGACCTTCGCTTT	GTATCCGAAGTGGCGG	540
Db	467	-----	-----	466
Qy	541	GTTATTGATGTCTCAGTTTCCACGAGGTTGAAAGACTTGAGAGCG	CGCCGGGGAAATGT	600
Db	467	-----GATGTGTCAGGTTCCACGAGTTGAAAGACTTGAGAGTG	CGCCGGGGAAATGT	520
Qy	601	CCTTTGTAGCCAGCTCAACGCTCAGCAGCTGTAAATGGCTGGGCG	CTCTCTACACGCC	660
Db	521	CCTTTGTAGCCAGCTCAGCGGTCAGCAGCTGTAATGGCTGGGGG	GTCTCTATGCTGCC	580
Qy	661	ACTGTGAAGAACTTCCCTGGGAGCTGACCCATCATCTCCGAGAG	CTGTGGGTGAGCTGAG	720
Db	581	ACTGTGAAAACTACCTTGGGACGAGGCCAAATTATCCAGAGACG	TGGTTCGTGCGCAG	640
Qy	721	GACTGGATTCGAACGAGACTTCTCATCTCGCTGGCTTAATGCTC	CGACGCTTTGTGCGACGT	780
Db	641	GACTGGATTCGACAGATACCTTCCTTCCTGGCTGAACGCCCGC	AGGCTTTGTGCGAGCC	700
Qy	781	ATGGTCTCTGAGCCAGCTGAGTGGGGGATGAAGATGGAGACGAT	GAATCTTTTTTTTTC	840
Db	701	GTGSCCTTGAGCCACGCGAATGSGGGATGAAGATGGAGACGAA	ATCTACTTCTTC	760
Qy	841	TTACGAGAGCTTCCGAGTGTGGACTTCCTATGAGGCGATCAAGT	TCCCAAGAGTGGCC	900
Db	761	TTTACGAGAGACTTCCCGAGCATTTTGACTCATACGAGCGCAT	TAAAGTCCACCGGTGGCC	820
Qy	901	CGAGTGTGTGGGGGACCTTGGGGGAGGAAGACCTTCAGCAGAG	ATCGACGACGCTTT	960
Db	821	CGTGTGTGTGGGGGACCTCGGGGGCCGAAGACCTCCAGCAGAG	ATGGACGACGCTTT	880
Qy	961	CTGAAGGCTGACCTGTGTGCCAGGCGCCGAGCATGGCGGGGCT	CCGGGGTCTGCAG	1020
Db	881	TTGAAGCTGACCTGTCTGTCCAGGSCCTGAGCATGSCCGGGCT	CCAGTCTCCAGTG	940
Qy	1021	GCTATGCCAGCTTGGGCTCAGCTTGAGCGGGGAACCCCATCT	TTTATGGGATCTTT	1080
Db	941	GATGTTCTGTGCTTCGACCTGAGCTTGGGGCAGGGACTCCCAT	CTTTTATGGCATCTTT	1000
Qy	1081	TCTCCCAAGTGGGAAGGAGCTGCATCTCTGCTGTGTGCGCT	TCGACCCCAAGACATC	1140
Db	1001	TCTTCCAGTGGGAGGGGGTACTATCTCTGCTGTGTGCTGCT	TCGACCCAGGACATTT	1060
Qy	1141	CGGCACTGTGAATGGTCCCTTTAGAGAGCTAAAACATGACTG	CAACAGGGGACTGCCT	1200
Db	1061	CGGACAGTGTGAATGTTCTTTCAGAGAACTAAACATGACTG	CAACAGGACTGCCT	1120
Qy	1201	GTCAATGACAACGAGGTGCCCGACGCCAGACCTGGAGAGTGCA	TGCGCAACACATGAAG	1260
Db	1121	GTGTTGACCAATGATGTGCCCGACGCCAGACCTGGAGAGTGCA	TCAACCAACACATGAAG	1180
Qy	1261	CTCAGCAGTTTGATCCTCACTCTCCCTGCCAGACCGCGTGT	CACTTTATCAGAGAC	1320
Db	1181	CTCGGCACTTTTGGCTCATCTCTCCCTGCCCTGACCGCGTAC	TCACTTATCCCGGAC	1240
Qy	1321	CAGCCTCTATGACAGCGCCGCTGTTCCCGGCTGAGGGCGCC	CGCTGCTGGTCACTACA	1380

[illegible]

Db 2318 ATCTAGAGCTGGGCAAAATGACCACCTAGTGTATATAAGTGATCACTGGAAC 2365

Search completed: September 26, 2003, 02:11:37
Job time : 14366.7 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2003, 19:18:34 ; Search time 7933.17 seconds
(without alignments)
12279.116 Million cell updates/sec

Title: US-09-284-180A-1

Perfect score: 4008

Sequence: 1 gccagggccgcgcagtagc.....aaaaaaaaaaaaaaaaaaaaa 4008

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1585.2	39.6	2719	11 AK035993	AK035993 Mus muscu
2	1582	39.5	2855	11 AK047067	AK047067 Mus muscu
3	645	16.1	704	14 CB527136	CB527136 UI-M-FY0-
4	633.8	15.8	657	14 CB557408	CB557408 AMGNNUC:N

5	632.6	15.8	691	12	BM944530	BM944530 UI-M-EH0P
6	612.2	15.3	868	10	BG323723	BG323723 602421820
7	610.2	15.2	701	14	CB245456	CB245456 UI-M-FY0-
8	596.6	14.9	1154	12	BM462308	BM462308 AGENCOURT
9	596.4	14.9	693	12	BM963828	BM963828 UI-M-EQ0-
10	586.2	14.6	901	13	BQ213659	BQ213659 AGENCOURT
11	573.8	14.3	765	12	BI823626	BI823626 603038515
12	566.2	14.1	758	14	CB244307	CB244307 UI-M-FY0-
13	557.6	13.9	733	10	AW953466	AW953466 EST365536
14	555.8	13.9	578	14	CB607655	CB607655 AMGNNUC:N
15	554.2	13.8	669	10	BB628682	BB628682
16	546.8	13.6	733	10	AW957134	AW957134 EST369204
17	509	12.7	520	14	CB715569	CB715569 AMGNNUC:N
18	508.8	12.7	772	12	BI915256	BI915256 603184795
19	495.2	12.4	769	13	BQ745165	BQ745165 UI-M-EH0P
20	475.6	11.9	578	12	BI344452	BI344452 372924 MA
21	475.2	11.9	669	12	BG774505	BG774505 602662571
22	472.2	11.8	482	10	BF392768	BF392768 UI-R-CA0-
23	455.6	11.4	486	10	BE982860	BE982860 UI-M-CG0P
24	441.8	11.0	466	12	BI275275	BI275275 UI-R-CX0-
25	437.2	10.9	466	9	AW123399	AW123399 UI-M-BH2.
26	424.2	10.6	434	10	BF393117	BF393117 UI-R-CA0-
27	418.2	10.4	689	10	BF467265	BF467265 UI-M-CG0P
28	409.8	10.2	418	10	BF392081	BF392081 UI-R-CA0-
29	405.2	10.1	434	9	AW123577	AW123577 UI-M-BH2.
30	402.2	10.0	616	14	BY731493	BY731493 182687 MA
31	397.4	9.9	556	10	BE683223	BE683223 BY290889
32	393	9.8	444	13	BY290889	BY290889 BY290889
33	388.4	9.7	1045	10	BF734978	BF734978 601570573
34	386.8	9.7	503	9	AA459837	AA459837 zx50g12.r
35	382	9.5	666	13	B0058887	B0058887 UI-M-FR0-
36	380	9.5	380	14	CB810579	CB810579 AMGNNUC:Y
37	379	9.5	749	12	BQ042411	BQ042411 UI-M-EMO-
38	375.2	9.4	522	10	BE667719	BE667719 155505 MA
39	370.8	9.3	379	10	BF392173	BF392173 UI-R-CA0-
40	366.4	9.1	368	10	BF415905	BF415905 UI-R-CA0-
41	362.8	9.1	731	12	BQ042997	BQ042997 UI-M-EG0P
42	352.2	8.8	416	13	BY224521	BY224521 BY224521
43	346.6	8.6	396	13	BY203844	BY203844 BY203844
44	319.8	8.0	425	13	BY237710	BY237710 BY237710
45	314.4	7.8	365	14	CD355327	CD355327 UI-M-FY0-

ALIGNMENTS

RESULT 1
AK035993
LOCUS
DEFINITION
AK035993 2719 bp mRNA linear HTC 05-DEC-2002
Mus musculus 16 days neonate cerebellum cDNA, RIKEN full-length
enriched library, clone:9630025H20 product:sema domain,
immunoglobulin domain (Ig), TM domain, and short cytoplasmic
domain, full insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AK035993
AK035993.1 GI:26084949
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374

QY	778	GCTATGGTCTGAGCCCGAGCTGAGTGGGGGATGAAGATGAGAGCATGAAATCTTTTT	837	DEFINITION	Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length
Db	816	GCTATGGTCTGAGCCCGCTGAGTGGGGGATGAAGATGAGAGCATGAAATCTTTT	875	ACCESSION	AK047067
QY	838	TTCTTCACGGAGACCTCCCGAGTGTGGACATCTATGAGCGCATCAAGGTCCCAAGAGTG	897	VERSION	AK047067.1 GI:26091917
Db	876	TTCTTCACGGAGACCTCCCGAGTGTGGACATCTATGAGCGCATCAAGGTCCCAAGAGTG	935	KEYWORDS	HTC; CAP trapper
QY	898	GCCCGAGTGTGTCGGGGGACCTTGGGGCAGGAAGACCCCTTCACGAGATGACGACG	957	SOURCE	Mus musculus (house mouse)
Db	936	GCCCGTGTGTCGGGGGACCTTGGGGCAGGAAGACCCCTTCACGAGATGACGACG	995	ORGANISM	Mus musculus
QY	958	TTTCTGAAGGTGACCTGCTGTGCCCGACGGGCCGAGCATGCGGGCTCCGGGTTCTG	1017	REFERENCE	1 Carninci, P. and Hayashizaki, Y.
Db	996	TTTCTAAAGGTGACCTGCTGTGCCCGACGGGCCGAGCATGGAAGGGCTCGGGGTTCTG	1055	AUTHORS	High-efficiency full-length cDNA cloning
QY	1018	CAGGCTATGGACGCTTCGGCCCTCAGCTCGAGCGGGACCCCTTCATTTATGGGATC	1077	JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
Db	1056	CAGGATATGACAGCTTCGACCTCAGCTCGCGGGGACCCCTTCATTTATGGGATC	1115	MEDLINE	99279253
QY	1078	TTTTTCTCCAGTGGGAAGGAGCTGCCATCTCTGTGTGCTTCGACCCCAAGAC	1137	PUBMED	10349636
Db	1116	TTTTTCTCCAGTGGGAAGGAGCCCATTTCTGTGTGCTTCGACCCCAAGAC	1175	REFERENCE	2 Carninci, P. and Hayashizaki, Y.
QY	1138	ATCCGGGACGTGCTGAATGGTCCCTTTTAGAGAGCTAAACATGACTCAACAGGGGACTG	1197	AUTHORS	Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Db	1176	ATCCGGGACGTGCTGAATGGTCCCTTTTAGAGAGCTAAACATGACTCAACAGGGGACTA	1235	TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to
QY	1198	CTGTGTCGACACAGGTGCCCCAGCCAGACCTGGAGAGTGCATCGCCCAACATG	1257	JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
Db	1236	CTGTGTCGACACAGGTGCCCCAGCCAGACCTGGAGAGTGCATCGCCCAACATG	1295	MEDLINE	20499374
QY	1258	AAGCTCCAGCAGTTTGGATCTCTACTCTCCCTGCGACACCGCTGCTTATCAGA	1317	PUBMED	11042159
Db	1296	AGTTCCAGCAGTTTGGATCTCTACTCTCCCTGCGACACCGCTGCTTATCAGA	1355	REFERENCE	3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
QY	1318	GACCACTCTCATGACAGGCCCGTGTTCGCCGTGACGGCCGCCCTGCTGGTCACT	1377	AUTHORS	Kunio, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
Db	1356	GACCACTCTCATGACAGGCCCGTGTTCGCCGTGACGGCCGCCCTGCTGGTCACT	1415	KEYWORDS	Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
QY	1378	ACAGATACAGCTATCTCAGAGTCGTGCCACAGGTGACCGCTCTCAGGAAAGAA	1437	SOURCE	Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Db	1416	ACGGATACAGCTATCTCAGAGTTGTGCTCACCAGGTGACCGCTCTCAGGAAAGAA	1475	ORGANISM	Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
QY	1438	TATGAGTCTCTACCTGGGACAGAGTGGACACCTCCACCGGGTGTGCGATTGGA	1497	REFERENCE	Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kasai, J.,
Db	1476	TATGAGTCTCTACCTGGGACAGAGTGGACACCTCCATCGGGTGTGCGATTGGA	1535	AUTHORS	Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
QY	1498	GTCAGCTCAGTGTCTTGAGAGATCGCCCTTGTTCAGAACACACCGGTTGAGAG	1557	TITLE	RIKEN integrated sequence analysis (RISA) system--384-format
Db	1536	GTCAGCTCAGTGTCTTGAGAGATCGCCCTTGTTCAGAACACACCGGTTGAGAG	1595	JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
QY	1558	ATGAATGTACACAGTATGGTCTGTGGGCTCCCATCTAGAGTGACACAGTGAAC	1617	MEDLINE	20530913
Db	1596	ATGAATGTACACAGTATGGTCTGTGGGCTCCCATCTAGAGTGACACAGTGAAC	1655	PUBMED	11076861
QY	1618	ACAGCAACTGTGGCGCTCCAGAGTGTCTGAGTGTATCTGCGCCAGACCCCGTG	1677	REFERENCE	4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Db	1656	ACAGCAACTGTGGCGCTCCAGAGTGTCTGAGTGTATCTGCGCCAGACCCCGTG	1715	AUTHORS	Akawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
QY	1678	TGCGCTGAGGCTTCGGGCTTGATCTTGTGTGGCCACCGCGGCGGAGTGATG	1737	KEYWORDS	Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Db	1716	TGCGCTGAGGCTTCGGGCTTGATCTTGTGTGGCCACCGCGGCGGAGTGATG	1775	SOURCE	Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R.,
QY	1738	GT 1739		ORGANISM	Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Db	1776	GT 1777		REFERENCE	Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
				AUTHORS	Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,
				TITLE	Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,
				JOURNAL	Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
				MEDLINE	Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
				PUBMED	Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
				REFERENCE	Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
				AUTHORS	Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
				KEYWORDS	Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
				SOURCE	Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
				ORGANISM	Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
				REFERENCE	Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilmings, L.,
				AUTHORS	Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
				TITLE	Functional annotation of a full-length mouse cDNA collection
				JOURNAL	Nature 409 (6821), 685-690 (2001)
				MEDLINE	21085660
				PUBMED	11217851
				REFERENCE	5 The FANTOM Consortium and the RIKEN Genome Exploration Research
				AUTHORS	Group Phase I & II team.

RESULT 2
AK047067
LOCUS

Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurinara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takada, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://phantom.gsc.riken.go.jp/.

FEATURES

Source

Location/Qualifiers

1..2855

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM DB:B930015M09"

/db_xref="taxon:10090"

/clone="B930015M09"

/tissue_type="cerebellum"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="10 days neonate"

1..2855

/note="sema domain, immunoglobulin domain (Ig), TM domain, and short cytoplasmic domain (MGD) MG1:1340055, GB|NM_011350, evidence: BLASTN, 99%, match=1662"

580 a 795 c 746 g 733 t 1 others

BASE COUNT

ORIGIN

Query Match 39.5%; Score 1582; DB 11; Length 2855;

Best Local Similarity 94.7%; Pred. No. 1.1e-175;

Matches 1649; Conservative 0; Mismatches 90; Indels 3; Gaps 1;

1 GCCAGAGCCGCGCAGTAGCGTACTAGTAGAGGCTGTGGACGCGCCGCCACCGCGGCAC 60

20 GCCAGAGCCGCGCAGTAGCGTACTAGTAGAGGCTGTGGACGCGCCGCCACCGCGGCAC 79

61 CAGGCGGAGCCAGAGATGTTGCCAGGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCT 120

80 CAGGCGGAGCCAGAGATGTTGCCAGGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCT 139

121 CGGCTCTTTCCTCCCGCGCGCGCTGTGCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 177

140 CCGGCTCTCTCTTCCCGCGCGCGCTGTGCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 199

178 GCCCGGGGTGTGGCGCGCGCTCCCGCGCTCAGTGCCGAGAACCTCGCTGCCATCTCCGAG 237

200 GCCCGGGGTGTGGCGCGCGCTCCCGCGCTCAGTGCCGAGAACCTCGCTGCCATCTCCGAG 259

238 GCTGACTCTTATCTACCGGGTTTGTAGGGTCTCATAGCTACAAATTAATCTGCTCTCTCT 297

260 GCTGACTCTTATCTACCGGGTTTGTAGGGTCTCATAGCTACAAATTAATCTGCTCTCTCT 319

298 GTGGATCTCTGCTCCCGAGAGCTTTACGTCGGTGCACGGGATAGCATCTTCGCTTAACC 357

320 GTGGATCTCTGCTCTCACACATTTACGTCGGTGCACGGGATAGCATCTTCGCTTAACC 379

358 CTCCTCTCTCTCTGGGGAAGACCCCGAAGGATTCGACTGGATGGTACCTCAGACTCACAGA 417

380 CTGCTCTCTCTCTGGGGAAGACCCCGAAGGATTCGACTGGATGGTACCTCAGACTCACAGA 439

418 CAGAACTGCAGGAAGAAAGCGAAGAGAGAGCAATGTACAAATTTTATCCAGATTCTC 477

440 CAGAAGCTGTAGGAAGAAAGCGAAGAGAGAGCAATGTACAAATTTTATCCAGATTCTC 499

478 GCATTTGTCATGTCCTCTACCTCTCTCAGCTGCGGCGACCTTCGCTTTTATCCGAGTGC 537

500 GCATTTGCAATGTCCTCTCAGCTGCGGCGACCTTCGCTTTTATCCGAGTGC 559

538 GGGGTATTGATGTGTCCAGTTTCCAGCAGGTTGAAGAGCTTGAGAGCGCGCGGGGAAA 597

560 GGGGTATTGATGTGTCCAGTTTCCAGCAGGTTGAAGAGCTTGAGAGTGGCGGGGAAA 619

598 TGTCTCTTTTGGCCAGCTCAACGGTCAGCAGCTGTAAATGGCTGGGGCGCTCTCTACAC 657

620 TGTCTCTTTTGGCCAGCTCAACGGTCAGCAGCTGTAAATGGCTGGGGCGCTCTCTACAC 679

658 GCATTTGTAAGAACTTCTCTGGGACTGAGCCCATCATCTCCGAGCTGTGGTCCAGCT 717

680 GCATTTGTAAGAACTTCTCTGGGACTGAGCCCATCATCTCCGAGCTGTGGTCCAGCT 739

718 GAGGACTGATTCGAACAGAGACCTTGTCTATCTGCTTAAATGCTCCAGCTTTGTGCGA 777

740 GAGGACTGATTCGAACAGAGACCTTGTCTATCTGCTTAAATGCTCCAGCTTTGTGCGA 799

778 GCTATGGTCTCAGCCAGCTGAGTGGGGGATGAAGATGGAGAGCATGAAATCTTTT 837

800 GCTATGGTCTCAGCCAGCTGAGTGGGGGATGAAGATGGAGAGCATGAAATCTTTT 859

838 TTCTTTCAGGAGACCTCCGAGCTGTGGACTCTTATGAGCGCATCAAGTCCCAAGAGTG 897

860 TTCTTTCAGGAGACCTCCGAGCTGTGGACTCTTATGAGCGCATCAAGTCCCAAGAGTG 919

898 GCCGAGTGTGTGCGGGGACCTTGGGGCAGGAGACCTTTCAGCAGAGATGGAGCAGC 957

920 GCCGAGTGTGTGCGGGGACCTTGGGGCAGGAGACCTTTCAGCAGAGATGGAGCAGC 979

958 TTCTTTCAGGAGTGTGCTGTGTCAGGCGCGGAGCAGTGGCGGCTTCGCGGGTCTTG 1017

980 TTCTTTCAGGAGTGTGCTGTGTCAGGCGCGGAGCAGTGGAGGCGCTTCGCGGGTCTTG 1039

1018 CAGGCTATGGCAGAGCTTCGGCTCAGCTGAGCGGGAGACCCCATCTTTTATGGATC 1077

1040 CAGGATATGACAGAGCTTCGACCTCAGCTGCGGGGAGACCCCATCTTTTATGGCATC 1099

1078 TTTTCTCTCCAGTGGGAGAGCTGCCATCTCTGCTGTGTGTGCTTCCGACCCCAAGAC 1137

1100 TTTTCTCTCCAGTGGGAGAGCGCCCATTTCTGCTGTGTGTGCTTCCGACCCCAAGAC 1159

1138 ATCCGGGAGTGTGAAATGGTCCCTTTAGAGAGCTTAAACATGACTGCAACAGGGAGCTG 1197

1160 ATCCGGGAGTGTGAAATGGTCCCTTTAGAGAGCTTAAACATGACTGCAACAGGGAGCTA 1219

1198 CCGTGTATGGCAACAGAGTGTGCCAGCCAGACCTTGAGAGTGTGATGCCAACACATG 1257

1220 CCGTGTATGGCAACAGAGTGTGCCAGCCAGACCTTGAGAGTGTGATGCCAACACATG 1279

1258 AAGCTCCAGCAGTTTGGATCTCTACTCTCCCTGCCAGACCGCTGCTACCTTTATCAGA 1317

1280 AAGTTCCAGCAGTTTGGATCTCTACTCTCCCTGCCAGACCGTGTGCTACCTTTATCAGA 1339

1318 GACCACTCTCATGGACAGCGCGTGTTCGGGCTGAGCGCGCGCGCGCTGCTGCTCACT 1377

1340 GACCACTCTCATGGACAGCGCGTGTTCGGGCTGAGTGGCGCGCGCGCTGCTGCTCACT 1399

1378 ACAGATACAGCCTATCTCAGAGTGTGGCCACAGAGGTGACAGGCTCTCAGGGAAGAA 1437

1400 ACGGATACAGCCTATCTCAGAGTGTGGCTACCGGGTGCACGCTCTCAGGGAAGAA 1459

1438 TATGAGTGTCTTACCTGGGAGAGAGGATGACACCTCCACCGGGCTGTGCGCATTTGA 1497

1460 TATGATGTCTTACCTGGGAGAGAGGATGGCACCTCCATCGGGCTGTGCGCATTCGGA 1519

QY 1498 GCTCAGCTCAGTCTGCTGGAGGATCTGGCCCTGTGTTCCAGAACACACAGCCGGTTGAGAGC 1557
 Db 1520 GCTCAGCTCAGTCTGCTGGAGGATCTGGCCCTGTGTTCCGGAACACACAGCCGGTTGAAAGC 1579
 QY 1558 ATGAAATTTGTACCACGATTTGGCTCCTGGTGGCTCCCATCTACTAGAGTGACACAGTGAAC 1617
 Db 1580 ATGAAATTTGTACCACGATTTGGCTCCTGGTGGCTCCCATCTACTAGAGTGACACAGTGAAC 1639
 QY 1618 ACCAGCAACTGTGGCCCTCTCCAGAGCTGCTCGGAGTGATPCTCGGCCAGGACCCCGTG 1677
 Db 1640 ACCAGCAACTGTGGCCCTCTCCAGAGCTGCTCGGAGTGATPCTCGGCCAGGATCCCGTG 1699
 QY 1678 TGCCTCGGAGCTTCCGGCTTGATGCTTGTGTGGCCACCGCCGCGAGACCCGCGGATG 1737
 Db 1700 TGTGCTCGGAGCTTCCGGCTCGATGCTTGTGTGGCCATGTCAGCGAGCACCGTGGGTGA 1759
 QY 1738 GT 1739
 Db 1760 GT 1761

RESULT 3
 CB527136
 LOCUS CB527136 704 bp mRNA linear EST 28-MAR-2003
 DEFINITION UI-M-FY0-cfh-i-24-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
 IMAGE: 6849601 5', mRNA sequence.
 CB527136
 VERSION
 KEYWORDS
 SOURCE EST.
 ORGANISM Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 704)
 NTH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.
 Location/Qualifiers
 1..704
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone_image="6849601"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_FY0"
 /note="Organ: Brain; Vector: pYX-Asc; Site_1: Ecor I;
 Site_2: Not I; The library was constructed according
 Ronald, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with Ecor I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is AGCAGACAG. This library was created for the University
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,

BASE COUNT 149 a 180 c 212 g 163 t
 ORIGIN
 Query Match 16.18; Score 645; DB 14; Length 704;
 Best Local Similarity 95.68; Pred. No. 3.3e-66;
 Matches 674; Conservative 0; Mismatches 30; Indels 1; Gaps 1;
 QY 380 CCCAAGGATCGACTGGATGCTGAGACTACAGACAGAACTGCAGGAAGGAAGCA 439
 Db 1 CTCGAGGATCGACTGGATGCTGCGCCGAGACTCAGACAGAACTGTAGGAAGGAAGCA 60
 QY 440 AGAAGAGGAGCAATGTCACAAATTTATCCAGATTCGCCCATTTGTCAATGCCCTCTCACC 499
 Db 61 AGAAGAGGAGCAATGTCACAAATTTATCCAGATTCGCCCATTTGTCAATGCCCTCTCACC 120
 QY 500 TCCTCAGCTGGCGGACCTTCGCTTTTGATCGAAGTGGGGGTTATTCATGTGCCAGTT 559
 Db 121 TCCTCAGCTGGCGGACCTTCGCTTTTGATCGAAGTGGGGGTTATTCATGTGCCAGTT 180
 QY 560 TCCAGCAGGTTGAAAGACTTGAGAGCGGCCGGGGAATGTCTTTTTCAGCAGCTCAAC 619
 Db 181 TCCAGCAGGTTGAAAGACTTGAGAGTGGCCGGGGAATGTCTTTTTCAGCAGCTCAAC 240
 QY 620 GGTGAGCAGCTGTAATGGCTGGGGCGCTCTCTACACGCCACCTGTGGAAGAACTTCCTGG 679
 Db 241 GGTGAGCAGCTGTAATGGCTGGGGCGCTCTCTACACGCCACCTGTGGAAGAACTTCCTGG 300
 QY 680 GGTGAGCAGCTGTAATGGCTGGGGCGCTCTCTACACGCCACCTGTGGAAGAACTTCCTGG 739
 Db 301 GGTGAGCAGCTGTAATGGCTGGGGCGCTCTCTACACGCCACCTGTGGAAGAACTTCCTGG 360
 QY 740 CTTTGTCTATCTGCTGCTTAATGCTCAGCCTTTGTCGAGCTATGCTGAGCCAGCTG 799
 Db 361 CTTTGTCTATCTGCTGCTTAATGCTCAGCCTTTGTCGAGCTATGCTGAGCCAGCTG 420
 QY 800 AGTGGGGGATGAAGATGGAGACGATGAAATCTTTTTTTTTCACGAGACCTCCCGAG 859
 Db 421 AGTGGGGGATGAAGATGGAGACGATGAAATCTTTTTTTTTCACGAGACCTCCCGAG 480
 QY 860 TGTGGGACTCTATGAGCGCATCAAGTCCCAAGAGTGGCCGAGTGTGGGGGAGC 919
 Db 481 TCTTGGACTCATATGAGCGCATCAAGTCCCAAGAGTGGCCGAGTGTGGGGGAGC 540
 QY 920 TTGGGGGAGGAGAACCCCTTCAGCAGAGATGGAGCAGCTTTCTGAAGGCTGACCTGCTGT 979
 Db 541 TTGGGGGAGGAGAACCCCTTCAGCAGAGATGGAGCAGCTTTCTGAAGGCTGACCTGCTGT 600
 QY 980 GCCCAGGCGCCGAGCATGGCCGGCTCCCGGGGTTCTCAGGCTATGCGAGAGCTTCGGC 1039
 Db 601 GTCCAGGCGCCGAGCATGGAAAGGCTCCGGGGTTCTCAGGATATGACAGAGCTTCGAC 660
 QY 1040 CTCAGCTGGAGCGGGAACCCCATCTTTTATGGATCTTTTCCT 1084
 Db 661 CTCAGCTGGAGCGGGAACCCCATCTTTTATGGATCTTTTCCT 704

RESULT 4
 LOCUS CB557408 657 bp mRNA linear EST 02-APR-2003
 DEFINITION AMGNNUC:NRHY3-00245-F12-A w Rat hypothalamus (10735) Rattus
 norvegicus cDNA clone nrhy3-00245-f12 5', mRNA sequence.
 CB557408
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE EST.
 ORGANISM Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 657)
 Amgen EST Program.
 Amgen Rat EST Program

BM944530
BM944530.1 GI:19428115
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 691)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi.
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 691)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project
(BNAP)

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project
(BNAP)
Seq primer: pVX-5.

This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pYX-5.
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

```
seq primer: pyx-5.
Location/Qualifiers
1. .691
```

```
1. 051
/organism="Mus musculus"
/mol_type="mRNA"
```

```

/db_xref="taxon:10090"
/strain="C57BL/6"

```

```
/clone="IMAGE:5695895"  
/tissue_type="whole brain"
```

```
/dev_stage="embryo 18.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/lab_name="lib-NTU BMAN EU02"
/lab_name="colony lib-NTU BMAN EU02"
```

```

/clonelib="NIH_BMAP_EHOp"
/note="Organ: brain; Vector: pYX-Asc; Site_1: EcoR I;
Site 2: Not I. The library was constructed according to

```

site-2. NOL 1, the library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded

CDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and the

cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CAGCCACAGC. This library was created for the

, is CAGCCACGAC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System'.

(bmt): Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

145 a	177 c	207 g	161 t	1 others
-------	-------	-------	-------	----------

15.8%; score 632.6; DB 12; Length 691;

Similarity	95.5%;	Pred. No.	9.2e-65;
61; Conservative	0;	Mismatches	30; Indels
			1; Gaps

3 CTGGATGGTACCTGAGACTCACAGACAGAACTGCAGGAGGAAGGCAAGAAAGAGGACGA 452

1 CTGGATGGTGTCCCGAGACTCACAGACAGAACTGTAGGAAGAAAGGCAAGAAAGAGGACGA 60

3 ATGTCAAAATTTATCCAGATTCTGCCATTGTCAATGCCCTCTCACCTCCTCACGTGGG 512

1 ATGTCACAAATTTTATCCAGATTCTCGCCATTGCCAATGCCTCTCACCTCCTCACGTTGG 120

33 CACCTTCGCTTTTATGATCCGAAGTCCGGGGTTATATGATGTGTCCAGTTTCCAGCAGGTTGA 572

1 I CACCTTCGGCTTTTGTATCCGAAGTCGGGGGTTATTGATGTGTCCAGTTTCCAGCAGGTTGA 180
3 AAGACTTCTCACACCCCGCCCGCGGGGGAATGTCCTCGGGGATGTCCTGACCCCGACAGCCCGGCG 532

1 AAGACTTGAGAGTGGCCGGGGGAAATGTCTCTTTTGGCCAGCTCAACGCTCAGCAGCTGT 240
 3 AAGACATGAGAGACGGCCCCGGGGGAAATGTCTCTTTGAGCCAGCTCAACCGGTCAGCAGCTGT 633

3 AATGGCTGGGGGGCGTCCCTCTACACGGCCACTGTGAAGAACATTCTCTGGGGACTGAGCCCAT 692

|||||
241 AATGGCTGGGGCGTCTCTACACCGCCACTGTGAATAACTTCTTGGGACAGAGCCGAT 300
QY |||||
693 CATCTCCGAGCTGTGGTGGAGCTGAGGACTCGAATTCGAACAGAGACCTTGTCACTCG 752
Db |||||
301 TATCTCCGAGCTGTGGTGGAGCTGAGGACTCGAATTCGAACAGAGACCTTGTCACTCG 360
QY |||||
753 GCTTAATGTCACGCTTTGTGCGAGCTATGTCCTGAGCCAGCTGAGTGGGGGATGA 812
Db |||||
361 GCTTAATGCCACGCTTTGTGCGAGCTATGTCCTGAGCCGCTGAGTGGGGGATGA 420
QY |||||
813 AGATGAGAGCATGAATCTTTTCTTTCACGGAGACCTCCGAGTGTGACACTCCTA 872
Db |||||
421 AGATGAGAGCATGAATCTTTTCTTTCACGGAGACCTCCGAGTGTGACACTCATA 480
QY |||||
873 TGAGCGCATCAAGGTCCCAAGAGTGGCCCGAGTGTGTGCGGGGAGCTTGGGGGACAGAA 932
Db |||||
481 TGAGCGCATCAAGGTCCCAAGAGTGGCCCGTGTGTGCGGGGAGCTTGGGGGCGGAA 540
QY |||||
933 GACCTTTCACGAGATGGACAGCTTTCTGAAAGGTGACCTGCTGTGCCAGGGCCGA 992
Db |||||
541 GACCTTTCACGAGATGGACAGCTTTCTTAAAGGCTGACCTGCTGTCCAGNCCGA 600
QY |||||
993 GCATGCGGGGCTCCGGGTTCTGAGGCTATGGGAGACCTTCGGGCTCAGCTGGAGC 1052
Db |||||
601 GCATGGAAGGCGCTCGGGGTTCTGAGGATATGACAGAGCTTCGACCTCAGCTGGCGC 660
QY |||||
1053 GGGAGCCCGC-TCTTTATGGGATCTTTTCT 1084
Db |||||
661 GGGAGCCCGC-TCTTTATGGGATCTTTTCT 691

RESULT 6

LOCUS BG323723
DEFINITION 602421820F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4560028 5',
mRNA sequence.
ACCESSION BG323723
VERSION BG323723.1 GI:13130160
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 868)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
plate: LLC1268 row: g column: 05
High quality sequence stop: 860.

FEATURES

source

1..868
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4560028"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_14"
/notes="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(kb). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in

the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
BASE COUNT 167 a 240 c 273 g 188 t

Query Match 15.3%; Score 612.2; DB 10; Length 868;
Best Local Similarity 83.4%; Pred. No. 1.9e-62;
Matches 721; Conservative 0; Mismatches 138; Indels 6; Gaps 2;
QY 1559 TGAATTTGTATACAGATTGGCTCCTGCTGGCGTCCCACTAGCTAGGTCACACAAGTGAACA 1618
Db 2 TGAATTTGTATACAGATTGGCTCCTGCTGGCGTCCCACTAGCTAGGTCACACAAGTGAATA 61
QY 1619 CCAGCAACTGTGGCCGCTCTCCAGAGCTGTCTCGAGTGTATCTTCGCGCCAGGACCCCGTGT 1678
Db 62 CAACCAACTGTGGCCGCTCTCCAGAGCTGTCTCGAGTGTATCTTCGCGCCAGGACCCAGTCT 121
QY 1679 GCSCCTGGAGCTTCCGCGTGTATGCTGTGTGGCCACGCGCGGGGAGCAGCGGGGATGG 1738
Db 122 GTGCCTGGAGCTTCCGCGTGGATGAGTGTGTGGCCATGTGCGCCATGCGGGGAGCAGCGGGTGG 181
QY 1739 TTCAAGATATAGAGTCAGCGGATGCTCTCTTCTTGTGTCACAAAGAACCTGGAGAACATC 1798
Db 182 TCCAAGACATAGAGTCAGCAGATGCTCTCTTCTTGTGTCCTAAAGAGCCTGGAGAACGTC 241
QY 1799 CCGTAGTGTGTTGAAGTTCGCGTGGCTACTGTGGGCCACGCTGGTCTGCCATGTTTCCCCCA 1858
Db 242 CAGTAGTGTGTTGAAGTTCGCGTGGCTACAGCTGGCATGTGGTCTGCCATGTTTCTCAA 301
QY 1859 GTTCTGCTGGGATCCTGTGTGTGGCACACGCGCCAGTGTGAGTGTACTCTACTCCCC 1918
Db 302 GCTCAGCATGGGATCCTGTGTGTGGCACACGCGCCAGTGTGAGTGTACTCTACCCCCC 361
QY 1919 GGAGGGATGACTAGAGTGTGTGTGACCCACGCGGCATGCGGGCTTATGCTTCCAGT 1978
Db 362 GCGGGGATGACTAGAGTGTGTGTGACCCACGCGGCATGCGGGCTTATGCTTGTGAAT 421
QY 1979 GTCAGGAGGTGGAGCGCGCGGTGTGTGTGTATAGCTTGTGTGGGCGAGCCAGC 2038
Db 422 GTCAGGAGGTGGGCGAGCCCATGTGTGTAGCAGCTTACAGCTTGGTATGGGCGAGCCAGC 481
QY 2039 GGGAGCCCTCAACCCGGGCGCCACACCGTGTGTGGGGGCTGAGTGTGGTGTCTCTGG 2098
Db 482 GAGATGCTCCGAGCGCGGGCGCCACACAGT--CGGGGGGCGGACTGGCTTGTCTTCTGG 539
QY 2099 GTCTTCTTGCAGCATCCT 2158
Db 540 GGATTTCTGCAGCATCCT 599
QY 2159 AGAGGGAGCTTCTAGCTAGAGACAAGGTGGGCTTAGATCTGGGGGTCTCCACCTTCTGGGA 2218
Db 600 AGAGGGAGCTTCTGGCTAGAGACAAGGTGGGCTTGGACCTGGGGCTTCCAACTTCTGGGA 659
QY 2219 CCACAAGCTATAGTACAGACCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2278
Db 660 CCACAAGCTACAGCAAGACCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 719
QY 2279 CCCTGGGTAAAGCGGGGCGAGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGT 2338
Db 720 CCCTGTCAAGAGGGGCGAGTGGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGT 779
QY 2339 GCCCAAGCCAGCCACATCCGCTCACTGGGGCGGCTCTAGCCACGCTGTGTATGAGACCT 2398
Db 780 GCCCAAGCCAGCCACATTCGGTAACCTGGGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 835
QY 2399 CCATCTAAAGCGGGGAAAAATGACT 2423
Db 836 CCATCTAAAGCGGGCAATGACCACT 860

RESULT 7

CB245456


```

/clone_lib="NIH_MGC_67"
/Note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
BASE COUNT      221 a      380 c      315 g      238 t
ORIGIN
Query Match      14.9%; Score 596.6; DB 12; Length 1154;
Best Local Similarity 82.9%; Pred. No. 1.1e-60;
Matches 699; Conservative 0; Mismatches 129; Indels 15; Gaps 1;
QY 1 GCCAGGCGCCGCGAGTAGGGTACTAAGTAGAGGCTGCTGAGCGCGCCACCCGCGAC 60
Db 1111111111111111111111111111111111111111111111111111111
Db 11 GCGGCGCAGTACGCCCGGGCCCTGAGCAGAGGCGGTAGCTTGCGCGCACCCGCGG 70
QY 61 CAGCGGAGCCAGAGATGCTGCCAGGCGCCAGCGGCCCGCGCGCGCGCGCGCT 120
Db 1111111111111111111111111111111111111111111111111111111
Db 71 CAGCGGAGCCAAAGATGCGGGCTCTGCTGCGCGGCCCGCGCGGTCCCGGCGAGCT 130
QY 121 CCGGTCTTTCCTTCCCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 1111111111111111111111111111111111111111111111111111111
Db 131 ACACCTCGCCCTCCC-----GCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 175
QY 181 CCGGTGTGCGCGCGCTGCCCGCTCAGTGCCCGAGAACCTCGCTGCCATCTCCGAGGCT 240
Db 1111111111111111111111111111111111111111111111111111111
Db 176 CCGGTATCCGCGCGCTCCCGCGCTCGGTGCCCGAGAACCTCGCTTCCAACTCTGAGGCT 235
QY 241 GACTCCTATCACCCTGTTGACGGCTCTCATAGCTACATTAATTAATTAATTAATTA 300
Db 1111111111111111111111111111111111111111111111111111111
Db 236 GACTCCTGTCTCACCCTGTTGCGAGTCCCTCAGCATACATTAATTAATTAATTAAT 295
QY 301 GATCCTGCTCCACACACTTTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 1111111111111111111111111111111111111111111111111111111
Db 296 GATCCTGCTCCACACACTTTATGTTGGCGCCCGGACACCACTTCCTGCTTATCCCTG 355
QY 361 CCTTCTCTGCGGGAAGACCCCGAGATGACATGATGATGATGATGATGATGATGATGAT 420
Db 1111111111111111111111111111111111111111111111111111111
Db 356 CCTTCTCAGGGAGAGACCCCGAGATTGACTGGATGTTCTTGAGGCTCACAGACAG 415
QY 421 AACTGCGAGGAAGACCAAGAGAGACCAATGTACAAATTTATCCAAATTTCTCGCC 480
Db 1111111111111111111111111111111111111111111111111111111
Db 416 AACTGTAGGAAGAAGCAAGAGAGACCAATGTACAAATTTGTCAGATTTCTCGCC 475
QY 481 ATTGTCAATGCTCTCACCTCTCACCTGCGCGCACCTTCGCTTTTGATCCGAAGTGGGG 540
Db 1111111111111111111111111111111111111111111111111111111
Db 476 ATTGCAATGCTCTCACCTCTCACCTGCGACCTTCGCTTTTGATCCGAAGTGGGG 535
QY 541 GTTATTGATGTCCAGTTTCCAGAGTTGAAAGACTTTGAGAGCGCGCGCGGGAATGT 600
Db 1111111111111111111111111111111111111111111111111111111
Db 536 GTTATTGATGTCCAGTTTCCAGAGTTGAAAGACTTTGAGAGTTGAGAGTTGAGAGTTG 595
QY 601 CCTTTGAGCCAGCTCAACGCTCAGCAGCTTAATGCTGGGGGCTCTACACGCC 660
Db 1111111111111111111111111111111111111111111111111111111
Db 596 CCTTTGAGCCAGCTCAGCAGCTTAATGCTGGGGGCTCTCTATGCTGCTGCC 655
QY 661 ACTGTGAAGAACTCTCTGGGACTGAGCCCATCATCTCCGAGCTGTGGGTGCGAGCTGAG 720
Db 1111111111111111111111111111111111111111111111111111111
Db 656 ACTGTGAAGAACTCTCTGGGAGCGGAGCCATTTATCAGCAGCAGTGGGTGCTGCCGAG 715
QY 721 GACTGGATTGCAACAGAGACCTTGTATCTCTGCTTAAATGCTCCAGCCTTTGTGCGAGCT 780
Db 1111111111111111111111111111111111111111111111111111111
Db 716 GACTGGATTGCGACAGATACCTTGCTTCTCTGCTGCTGAACGCGCCCGAGCCTTTGTG 775
QY 781 ATGCTCTGAGCCAGCTGAGTGGGGGATGAAAGATGAGAGCATGAATCTTTTTC 840
Db 1111111111111111111111111111111111111111111111111111111
Db 776 GTGGCTTTGAGCCCGAGGAAATGGGGGATGAAAGATGGAAGACGACCAATCTACTTC 835
QY 841 TTC 843
Db 836 TTC 838
```

```

RESULT 9
BM963828      693 bp      mRNA      linear      EST 18-MAR-2002
LOCUS      UI-M-EQ0-bwn-j-03-0-UI-r1.NIH_BMAP_EQ0 Mus musculus CDNA clone
DEFINITION      IMAGE:5698658 5', mRNA sequence.
ACCESSION      BM963828
VERSION      BM963828.1 GI:19547248
KEYWORDS      EST.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: crapsb-r@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5
Location/Qualifiers
1. .693
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5698658"
/tissue_type="whole brain"
/dev_stage="embryo 15.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_EQ0"
/Note="Organ: brain; Vector: pYX-Asc; Site_1: EcoR I;
Site_2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
, is CAGCCACGAC. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP). 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institutes of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."
BASE COUNT      124 a      180 c      231 g      157 t      1 others
ORIGIN
Query Match      14.9%; Score 596.4; DB 12; Length 693;
Best Local Similarity 93.1%; Pred. No. 1.5e-60;
Matches 646; Conservative 0; Mismatches 46; Indels 2; Gaps 2;
QY 1546 CCGGTTGAGACATGAATTTGTACACGATTTGGCTCTCTGGTGGCTCCCTACTAGAGTG 1605
Db 1111111111111111111111111111111111111111111111111111111
Db 1 CCGGTTGAACCATGAATTTGTACCATGATTTCTGGTGGCT-CCATCATGAGGTG 59
QY 1606 ACACAAGTGAACACCAAGCACTGTGGCGCTCTCAGAGCTGCTGGAGATGTATCTGCC 1665
Db 60 ACACAAGTGAACACCAAGCACTGTGGCGCTCTCCAGAGCTGCTAGAATGTATCTGCC 119
QY 1666 CAGGACCCGCTGTGGCGCTGAGCTTCGGCTTGATGCTTGTGTGGCCACCGCGCAG 1725
Db 120 CAGGATCCCGTGTGTGGCTGAGCTTCGGCTCGATGCTTGTGTGGCCCGCAGCGAG 179
```

QY 1726 CACCGGGGATGGTTCAAGATATAGAGTCAGCGGATGCTCTCTTTGTGTCACAAAGAA 1785
Db 180 CACCGTGGGATGGTTCAAGACATAGAGTCAGAGATGCTCTCTTTGTGTCACAAAGAA 239
QY 1786 CCTGAGAGACATCCCGTAGTGTTCAGATTCGGGTGGCTACTGTGGCCACGTGTCCTG 1845
Db 240 CCTGAGAGACATCCCGTAGTGTTCAGATTCAGTGGCTAGCTGGGTCACGTGTCCTG 299
QY 1846 CCATGTTCCCGCAGTTCCTCGCTGGGCATCTGTGTGGCCACACCGAGTGGAGTGACT 1905
Db 300 CGGTGCTCCCGCAGTTCCTCGCTGGGCATCTGTGTGGCCACACCGAGTGGAGTGACT 359
QY 1906 GGGCTACTCCCGGAGGATGAGTATAGAGTGGTGGTGCACCCAGGGGCCATGGGGCT 1965
Db 360 TCGCTCAGCCACGAAAGGATGGGCTAGAGTGGTGGTGCACCCAGGGGCCATGGGGCT 419
QY 1966 TATGCTTCGGAGTGCAGAGGGTGGAGCGCCCGCTGGTGGCTTATAGCTTGGTG 2025
Db 420 TATGCTTCGGAGTGCAGAGGGTGGAGCGCCCGCTGGTGGCTTATAGCTTGGTG 479
QY 2026 TGGGCGAGCCAGCGGGGACCTCAAAACCGGGGCCACACCGCTGTG-GGGGGCTGGATTGGT 2084
Db 480 TGGGCGAGCCAGAGGGGCCCGCAAAACCGGGGCCACACCGTGTGNGGGGGCTGGGCTGGT 539
QY 2085 TGGCTTCTCTGGGTGTCTTTCAGCATCTCTACTCTCTCTCTGATGGTGGCCGTCA 2144
Db 540 TGGCTTCTCTGGGTGTCTTTCAGCATCTCTACTCTCTCTCTGATGGTGGCCGTCA 599
QY 2145 GCAGCGTGGCGACAGAGGGAGCTTCTAGCTAGACAGAGTGGGCTTATAGCTGGGGC 2204
Db 600 GCAGCGGCGCGACAGAGGGAGCTTCTAGCTAGACAGAGTGGGCTTATAGCTGGGGC 659
QY 2205 TCCACCTTCTGGGACCAAGCTATAGTCAGGAC 2238
Db 660 CTCACCTTCTGGGACCAAGCTATAGTCAGGAC 693

RESULT 10

BQ213659
LOCUS BQ213659 901 bp mRNA linear EST 02-MAY-2002
DEFINITION AGENCOURT_7595087 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6062683
5' mRNA sequence.
ACCESSION BQ213659
VERSION BQ213659.1 GI:20395053
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapb-r@mail.nih.gov
Tissue Procurement: ATCC/DC/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM1334 row: 1 column: 20
High quality sequence stop: 604.

FEATURES

source

RESULT 11
B1823626
LOCUS B1823626 765 bp mRNA linear EST 04-OCT-2001
DEFINITION 603038515F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5179587 5',
mRNA sequence.
ACCESSION B1823626
VERSION B1823626.1 GI:15935176
KEYWORDS EST.
SOURCE Homo sapiens (human)

/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies".

BASE COUNT	181 a	248 c	289 g	183 t
ORIGIN				
Query Match	14.6%	Score 586.2	DB 13	Length 901
Best Local Similarity	86.3%	Pred. No. 2e-59		
Matches 648	Conservative 0	Mismatches 103	Indels 0	Gaps 0
QY 1358	GC	GC	GC	GC
Db 12	GC	GC	GC	GC
QY 1418	CC	CC	CC	CC
Db 72	CC	CC	CC	CC
QY 1478	AC	AC	AC	AC
Db 132	AC	AC	AC	AC
QY 1538	AA	AA	AA	AA
Db 192	AA	AA	AA	AA
QY 1598	CT	CT	CT	CT
Db 252	CT	CT	CT	CT
QY 1658	TC	TC	TC	TC
Db 312	TC	TC	TC	TC
QY 1718	CG	CG	CG	CG
Db 372	CG	CG	CG	CG
QY 1778	CA	CA	CA	CA
Db 432	CA	CA	CA	CA
QY 1838	TG	TG	TG	TG
Db 492	TG	TG	TG	TG
QY 1898	GA	GA	GA	GA
Db 552	GA	GA	GA	GA
QY 1958	TG	TG	TG	TG
Db 612	TG	TG	TG	TG
QY 2018	GT	GT	GT	GT
Db 672	GT	GT	GT	GT
QY 2078	GA	GA	GA	GA
Db 732	GA	GA	GA	GA

```

ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaaps-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1447 Row: n Column: 04
High quality sequence start: 3
High quality sequence stop: 765.
Location/Qualifiers
1..765
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5179587"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector:
PCMV-SPO16; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: This is a NIH_MGC Library."
BASE COUNT 139 a 240 c 219 g 167 t
ORIGIN
Query Match 14.38; Score 573.8; DB 12; Length 765;
Best Local Similarity 85.58; Pred. No. 6.1e-58;
Matches 668; Conservative 0; Mismatches 97; Indels 16; Gaps 2;
Qy 26 TAAGTAGAGGCTGCTGGACGCGCCGCCACCGGACCGGAGCGGAGATGCTTGCCA 85
Dy 1 TGACAGAGGCGGTAGCTTGGCGCGACCCGCGGAGCGGAGCGGAGATGCTTGCCA 60
Qy 86 GGGCCGAGCGCCCGCCGCGGCGCCCGCGGCGCTTCCGCTGCTTCCCTTCCGCGCGC 145
Dy 61 CTGCTGGCGG-CGCGCGCGGCTCCGCGGAGCGCTACAGCGCTCGCCCTTCC----- 111
Qy 146 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 205
Dy 112 -----GCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 164
Qy 206 CAGTGCCGAGAACCTCGCTGCCATCTCCGAGGCTGCTCTATCTACCCCGGTTTGAG 265
Dy 165 CGGTGCCGAGAACCTCGCTGCCATCTCTGAGGCTGCTCTGCTCTACCCGTTGCGAG 224
Qy 266 CGTCTCATAGCTACATTAATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 325
Dy 225 TCCCTCACACATAAATTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTATG 284
Qy 326 TCGTGTGACGGGATAGATCTTCTGCTTTTAACTTCTCTCTCTCTCTCTCTCTCT 385
Dy 285 TTGCGCGCGCGGAGACCATCTTCTGCTTTTATCCCTTCTCTCTCTCTCTCTCTCT 344
Qy 386 GGATCGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 445
Dy 345 GGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 404
Qy 446 AGGACGAATGTACAAATTTTATCCAGATTTCTCGCCATTGTCAATGCCTCTCACCTCTCA 505

```

```

Db 405 AGGACGAATGTACAAATTTTGTCCAGATTCTCGCCATTGCCAATGCCTCTCACCTCTCA 464
Qy 506 CGTCCGCGACCTTCGCTTTTTCATCCGAGTCCGGGGTATTGATGTGCTCCAGTTCCAGC 565
Dy 465 TTTTGTGCGACCTTCGCTTTTTCATCCGAGTCCGGGGTATTGATGTGCTCCAGTTCCAGC 524
Qy 566 AGGTTGAAGACTTTGAGAGCGCGCGGGAATATCTCTTTTTCGAGCCAGCTCAACGGTCTCAG 625
Dy 525 AGGTTGAAGACTTTGAGAGTGGCGCGGGAATATCTCTTTTTCGAGCCAGCTCAGCGGTCTCAG 584
Qy 626 CAGCTGTAATGGCTGGGGGCTCTCTACACCGGACCTGTGAGAACCTTCCTGGGGAGCTG 685
Dy 585 CAGCTGTAATGGCTGGGGGCTCTCTATGCTGCACCTGTGAAAAAATACCTGGGGAGCGG 644
Qy 686 AGCCCATCATCTCCGAGCTGTGGTGCAGCTGAGGACTGGATTTCGAACAGACACCTTGT 745
Dy 645 AGCCATATATCACAGACAGTGGTGTGTCGAGGAGCTGGATTTCGAGACATACCTTTCG 704
Qy 746 CATCTGGCTTAATGCTCCAGCTTTGTTCGAGCTATGGTCTCTGAGCCAGCTGAGTGGG 805
Dy 705 CTTCCTGGCTGAACGCGCCAGCTTTGTTCGAGCGGTGGCTTTCGAGCCAGCCCAATGGG 764
Qy 806 G 806
Dy 765 G 765

RESULT 12
CB244307
LOCUS
DEFINITION CB244307 758 bp mRNA linear EST 12-FEB-2003
IMAGE: 6832579 5', mRNA sequence.
ACCESSION CB244307
VERSION CB244307.1 GI:28365951
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 758)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaaps-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
The following repetitive elements were found in this cDNA
sequence: 91-124, >C-rich#Low complexity 141-182, >(CAG
)n#Simple_repeat (matched complement)
Seq primer: pfx-5.
Location/Qualifiers
1..758
/organism="Mus musculus"
/mol_type="mRNA"
/strain="c57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6832579"
/tissue_type="whole brain"
/dev_stage="embryo 13.5, 14.5, 16.5, 17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_MGC_FY0"
/note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;
Site_2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

```


1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

134 a 245 c 216 g 163 t

Query Match 14.1%; Score 566.2; DB 14; Length 758;
Best Local Similarity 83.9%; Pred. No. 4.7e-57;
Matches 706; Conservative 0; Mismatches 38; Indels 97; Gaps 2;

QY 1 GCCGAGCCCGCAGTAGTCTTAAGTAGAGGCTGCTGGAGCGGCCCGCCACCGCGCAC 60
|||||
Db 6 GCCGAGCCCGCAGTAGTCTTAAGTAGAGGCTGCCCGACGCGGCCCGCCACCGCGCAC 65

QY 61 CAGCGGAGCCAGAGATGCTTGCAGAGCCGAGGCGGCCCGCGGGCGCGCGCT 120
|||||
Db 66 CAGCGGAGCCAGAGATGCTGCGCAGGCGCGAGGCGGCCCGCGGGCGCGCGCT 125

QY 121 CCGGTCTTCTCCCGCGCGCGCTGTC---GTCGTGCTGCTGCTGGCGCATTAAGC 177
|||||
Db 126 CCGGTCTCTCTCTCCCGCGCGCGCTGTCGTGCTGCTGCTGCTGGCGCATGCTGAGC 185

QY 178 GCCCGGCTGTGGCGCGCGTCCCGCGCTCAGTGTGCCAGAACCTGCTGCCATCTCCGAG 237
|||||
Db 186 GCCCGGCTGTGGCGCGCGTCCCGCGCTCGTGTGCCAGAACCTGCTGCCATCTCCGAG 245

QY 238 GCTGACTCTATCTACCCGGTTTGCAGCGTCTCATAGCTACAATTACTGCTCTCGCTT 297
|||||
Db 246 GCTGACTCTATCTACCCGGTTTGCAGCGTCTCATAGCTACAATTACTGCTCTCGCTG 305

QY 298 GTGGATCTGCTCTCCACACACTTTACGTGCTGTGCAGGGATAGCATCTTCGCTTTAAC 357
|||||
Db 306 GTGGATCTGCTCTCCACACACTTTACGTGCTGTGCAGGGATAGCATCTTCGCTTTAAC 365

QY 358 CTCGCCCTTCTCTGGGGAAGACCCGAGAGTGCAGTGTGATGCTGATGCTGAGACTCACAGA 417
|||||
Db 366 CTCGCCCTTCTCTGGGGAAGACCCGAGAGTGCAGTGTGATGCTGATGCTGAGACTCACAGA 425

QY 418 CAGAACTGCAGAGAAAGGCAAGAAAGAGGACGAAATGTCACAATTTTATCCAGATTCTC 477
|||||
Db 426 CAGAACTGTAGAAAGGCAAGAAAGAGGACGAAATGTCACAATTTTATCCAGATTCTC 485

QY 478 GCCATTGTCAATGCTCTCAGCTCCTCAGTGGCGACCTTCGCTTTTGTATCCGAAGTGC 537
|||||
Db 486 GCCATTGTCAATGCTCTCAGCTCCTCAGTGTGGCAGCTTCGCTTTTGTATCCGAAGTGC 545

QY 538 GGGGTATTGATGTGTCAGTTTCCAGCAGGTTGAAAGACTTGAGAGCGCGCGGGGAAA 597
|||||
Db 546 GGGGTATT----- 554

QY 598 TGTCTTTTGTAGCCAGCTCAACGGTCAGCAGCTGTAATGGCTGGGGGGTCTCTACACC 657
|||||
Db 555 -----GGGGGCTCTCTACACC 571

QY 658 GCCACTGTGAAGAACTTCTGGGACTGAGCCCATCTCCCGAGCTGTGGGTCGAGCT 717
|||||
Db 572 CCCACTGTGAAGAACTTCTGGGACAGACCGGATATCTCTCGAGCTGTGGGTCGAGCT 631

QY 718 GAGGACTGGATTGCAACAGAGACCTTGTATCTGCTGCTTAATGCTCCAGCCTTTGTGCA 777
|||||
Db 632 GAGGACTGGATTGCAACAGAGACCTTGTATCTGCTGCTTAATGCTCCAGCCTTTGTGCA 691

QY 778 GCTATGGTCTTGAGCCCGAGCTGAGTGGGGGATGAAGATGGAGAGCAATCTTTT 837
|||||

Db 692 GCTATGCTCTAGAGCCCGGCTGAGTGGGGGATGAAGATGAGAGCATGAATCTCTTTT 751

QY 838 T 838
|
Db 752 T 752

RESULT 13
AW953466
LOCUS AW953466 733 bp mRNA linear EST 01-JUN-2000
DEFINITION EST1365536 MAGE resequences, MAGB Homo sapiens cDNA, mRNA sequence.
ACCESSION AW953466
VERSION AW953466.1 GI:8143149
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 733)
AUTHORS Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gasparid,R., Gay,C., Holt
I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
Quackenbush,J.
TITLE Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished
COMMENT The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@tigr.org
Plate: 51
Seq primer: Reverse.
Location/Qualifiers
1...733
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGB"
/note="vector: pBluescriptSKm"
BASE COUNT 133 a 207 c 238 g 154 t 1 others
ORIGIN
Query Match 13.9%; Score 557.6; DB 10; Length 733;
Best Local Similarity 86.5%; Pred. No. 4.8e-56;
Matches 627; Conservative 0; Mismatches 95; Indels 3; Gaps 1;

QY 1560 GAAATTGTACCACGATTTGGCTCTCTGGTGGGCTCCCATCTAGTGTACACAGTGAACAC 1619
|||||
Db 11 GAAATTGTACCACAGCTGGCTCTCTGGTGGGCTCCCATCTAGTGTACACAGTGAATAC 70

QY 1620 CAGCAACTGTGGCGCTCTCCAGAGCTGCTCGGAGTGTATCTGGCCCGAGGACCCGCTGTG 1679
|||||
Db 71 AACCAACTGTGGCGCTCTCCAGAGCTGCTCAGAGTGTATCTGGCCCGAGGACCCAGTGTG 130

QY 1680 CGCCTGGAGCTTCCGGCTTGATGCTTGTGTGCCCGACCGCGCGAGCACCAGGAGTGGT 1739
|||||
Db 131 TGCCTGGAGCTTCCGGCTGGATGATGTGTGGCCCATCGCGGGGAGCACCAGGCTGGT 190

QY 1740 TCAAGATATAGACTCAGCGGATGCTCTCTTTGTGTCCAAAGAAACCTGGAGACATCC 1799
|||||
Db 191 CCAGACATAGAGTACAGAGATGCTCTCTTTGTGTCTTAAGAGCCTGGAGACGCTCC 250

QY 1800 CGTAGTGTTCAGCTTCCGGTGGCTACTGTGGGCGACGCTGCTCTCCATGCTTCCCCCAG 1859
|||||
Db 251 AGTAGTGTTCAGCTTCCGGTGGCTACTGCTGCGCATGTGTGCTTGCATGCTTCTCCAG 310

QY 1860 TTCTGCTGGGCATCTGTGTGTGGCACACAGCCAGTGGAGTGTGCTGCGCTCACTCCCG 1919
|||||
Db 311 CTCAGCATGGGCATCTGTGTGTGGCACACAGCCAGTGGAGTGTGCTGCACTCACCCCG 370

QY 1920 GAGGATGGAGTACAGTGGTGGTGGCCCGAGGCGGCTTATGCTTGGAGTGT 1979
|||||

```
Db 371 GCGGGATGGACTGGAGGTGTGTGTACCCAGGGGCGATGGGCGCTTATGCCCTGTGAATG 430
Qy 1980 TCAGGAGGTTGGAGCCCGCCGCTGTGTGCTGTATAGCTTGTGTGGGCGACCCAGCG 2039
Db 431 TCAGGAGGTTGGGCGACCCATGTGTAGCAGCTTACAGCTTGTGTGGGCGACCCAGCG 490
Qy 2040 GGGACCCCTCAAAACGGGCGCCACACCGTTGTGGGGGCTGGATGTGTGGCTTCTCCTGGG 2099
Db 491 AGATGCTCCGAGCGCGGCGCCACAC- --AGTGGGGGGGCGACTGCTGCTTCTTCTTGGG 547
Qy 2100 TGTTCCTGACAGTCCCTCACTCTCTCTCTGATTTGTCGCGCTCAGCAGCGTGGCGACA 2159
Db 548 GATTCGTCAGCATCCCTGACTCTCAATTCCTGATTTGTCGCGCTCAGCAGCGAGCGGACA 607
Qy 2160 GAGGAGCTTCTAGCTAGACACAAGTGGGCTTAGACTCTGGGGCTCCACCTTCTGGGAC 2219
Db 608 GAGGAACTTCTGCTANAGACAAGTGGGCTTGGACCTGGGGCTCCACCTTCTGGGAC 667
Qy 2220 CACAAGCTATAGTCAGGACCCCTCCCTCTCTCTGCTTCCCTTGAAGATGAAGCGCTGCCCTGGC 2279
Db 668 CACAAGCTACAGGCAAGACCCCTGCTCCCTCTCTCTGAAAGATGAGCGGTTGCCCTGGC 727
Qy 2280 CCTGG 2284
Db 728 CCTGG 732

RESULT 14
CB607655
LOCUS CB607655 578 bp mRNA linear EST 16-MAY-2003
DEFINITION AMGNMUC:NRH5-00021-G9-A W Rat hypothalamus (10471) Rattus
norvegicus cDNA clone nrh5-00021-g9 5', mRNA sequence.
ACCESSION CB607655
VERSION CB607655.1 GI:29547268
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 578)
AUTHORS Angen EST Program.
TITLE Angen Rat EST Program
JOURNAL Unpublished
COMMENT Contact: Dan Fitzpatrick
Angen, Inc
One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00021 row: g column: 9.
FEATURES
source
1. 578
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="nrh5-00021-g9"
/notes="vector: pSPORT1; Site 1: SalI; Site 2: NotI; W Rat
hypothalamus adult female Wistar rat avg. insert size 2.3
kb fraction 6 and 7"
BASE COUNT 100 a 160 c 174 g 144 t
ORIGIN

Query Match 13.9%; Score 555.8; DB 14; Length 578;
Best Local Similarity 99.5%; Pred. No. 8.9e-56;
Matches 568; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
Qy 2553 TTCTGATTCATGAGATCAACCTGTAACTTCTGCGATGGCTTGTGTGGG 2612
Db 9 TCCGTGATTCATGAGATCAACCTGTAACTTCTGCGATGGCTTGTGTGGG 68
Qy 2613 CCATCAGCTTGTGGGTGGAGTAGGACATAGGCCCGGAAAGGAATCAGTGGAGGT 2672
Db 69 CCATCAGCTTGTGGGTGGAGTAGGACATAGGCCCGGAAAGGAATCAGTGGAGGT 128
```

```
Qy 2673 AGTTGGGGCGTGTGTCCTTGCCTTGTGTGGTGTATGATTTCCCACTGCTGAC 2732
Db 129 AGTTGGGGCGTGTGTCCTTGCCTTGTGTGGTGTATGATTTCCCACTGCTGAC 188
Qy 2733 TCTGGGAGCGCATGATCCCTTGCCTTGTGTAGATCTCCCAACTCAGTTTCCCTTG 2792
Db 189 TCTGGGAGCGCATGATCCCTTGCCTTGTGTAGATCTCCCAACTCAGTTTCCCTTG 248
Qy 2793 CTCTGGAAGAGTGTGTCTATACACTGTGTGCTTGAAGAGGCTGTCCATGTGTGATG 2852
Db 249 CTCTGGAAGAGTGTGTCTATACACTGTGTGCTTGAAGAGGCTGTCCATGTGTGATG 308
Qy 2853 GAGCAGACGGCGGCTCGCTTGTGGGAGTCGGAGAGAAAGTTTGAATGGGG 2912
Db 309 GAGCAGACGGCGGCTCGCTTGTGGGAGTCGGAGAGAAAGTTTGAATGGGG 368
Qy 2913 ACAACTTAACCTCGGTAGCAGTGAAGGAAACACATGCCCTGCCCATCACCCACAG 2972
Db 369 ACAACTTAACCTCGGTAGCAGTGAAGGAAACACATGCCCTGCCCATCACCCACAG 428
Qy 2973 CGCTTCTTAACCTTGAAGAAAGTTCCCAAGTGACCTTCTGGTGGGAGGCGAGG 3032
Db 429 CGCTTCTTAACCTTGAAGAAAGTTCCCAAGTGACCTTCTGGTGGGAGGCGAGG 488
Qy 3033 ACATGTGGCGCGCTCTCTCTTGTCTTTCCTTCTGGCTGCACCACTGCCGTGCC 3092
Db 489 ACATGTGGCGCGCTCTCTCTTGTCTTTCCTTCTGGCTGCC-ACCAGTCCGTGCC 547
Qy 3093 ACCGCTCGCTTTCCTTGGCTGAGTGGAGG 3123
Db 548 ACCGCTCGCTTTCCTTGGCTGAGTGGAGG 578
```

```
RESULT 15
BB628682
LOCUS BB628682 669 bp mRNA linear EST 26-OCT-2001
DEFINITION MUSCULUS CDNA clone 9630025H20 5', mRNA sequence.
ACCESSION BB628682
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 669)
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toyota,T.,
Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp/
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura
,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
```

Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa
K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
Hayashizaki, Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.

FEATURES
source
1. .669
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="9630025H20"
/tissue_type="cerebellum"
/dev_stage="16 days neonate"
/lab_host="DH10B"
/clone.lib="RIKEN full-length enriched, 16 days neonate
cerebellum"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTCTTTN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 20.0 and subtraction to Rot = 370.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGAGATCTCGAGTTAATTAAATATCCGCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I."
BASE COUNT 119 a 220 c 194 g 136 t
ORIGIN

Query Match 13.8%; Score 554.2; DB 10; Length 669;
Best Local Similarity 94.8%; Pred. No. 1.3e-55;
Matches 606; Conservative 0; Mismatches 28; Indels 5; Gaps 3;
QY 1 GCCGAGGCCGCGCAGTAGCGGTACTAGTAGAGGCTGCTGGACGCGCCGCCACCGGCAC 60
|||||
DB 33 GCCGAGGCCGCGCAGTAGCGGTACTAGTAGAGGCTGCGCGCGCCGCCACCGGCAC 92
QY 61 CAGCGGAGCCAGAGATGCTTTCAGAGGCCGAGCGGCCGCCGCCGCCGCCGCCGCCGCC 120
|||||
DB 93 CAGCGGAGCCAGAGATGCTGCGCCAGCGGCCGAGCGGCCGCCGCCGCCGCCGCCGCC 152
QY 121 CCGGTCTTTCCTCCCGCCCGCGCTGTC---GCTGCTGCTGCTGCTGGCGATACTAAGC 177
|||||
DB 153 CCGGTCTCTCTCTTCCCGCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 212
QY 178 GCCCGGTGTGCGCGCGCGTCCCGCTCAGTCCGAGAGCTCGCTGCCATCTCCGAG 237
|||||
DB 213 GCCCGGTGTGCGCGCGGTCCCGCTCAGTCCGAGAGCTCGCTGCCATCTCCGAG 272
QY 238 GCTGACTCTATCTACCCGGTTTGCAGCGCTCTCATACGTACAATTACTTGTCTCTCTT 297
|||||
DB 273 GCTGACTCTATCTACCCGGTTTGCAGCGCTCTCATACGTACAATTACTTGTCTCTCTG 332
QY 298 GTGGATCCTGCCTCCACACACTTTTACGTGGTGCAGGGGATAGCATCTTGGCTTTAAC 357

Db 333 GTGGATCCTGCTTCTCACACACTTTTACGTCGGTGCACGGGACAGCATCTTCGCTTTAAC 392
QY 358 CTCCCTTCTCTGGGGAAGACCCCGAAGGATCGACTCGATGGTACCTGAGACTCACAGA 417
Db 393 CTGCCCTTCTCTGGGGAAGACCCCGAAGGATCGACTCGATGGTACCTGAGACTCACAGA 452
QY 418 CAGAACTGCAGGAAGAGGCAAGAAAGAGGACGAATGTCAAAATTTTATCCAGATTCTC 477
Db 453 CAGAACTGTAGGAAGAGGCAAGAAAGAGGACGAATGTCAAAATTTTATCCAGATTCTC 512
QY 478 GCCATTGTCAATGCTCTCACCTCCTCACGTGCGGACCTTTCGCTTTTGTATCCGAAGTGC 537
Db 513 GCCATTGTCAATGCTCTCACCTCCTCACGTGCGGACCTTTCGCTTTTGTATCCGAAGTGC 572
QY 538 GGGGTTATTGATGTGTCCAGTTTCCAGCAGCTTGAAGACTTGAGAGCGCGCGGGGAAA 597
Db 573 GGGGTTATTGATGTGTCCAGTTTCCAGCAGCTTGAAGACTTGAGAGTGGCC-GGGGAAA 631
QY 598 TGTCTTTTGGAGCCAGCTCAACGGTTCAGCAGCTGTAAATG 636
Db 632 TGTCTTTTGGAGCCAGCTCAAC-GTCAGCAGCTGTAAATG 669

Search completed: September 26, 2003, 05:40:45
Job time : 7943.17 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 26, 2003, 02:11:50 ; Search time 620.263 seconds
(without alignments)
8552.073 Million cell updates/sec

Title: US-09-284-180A-1
Perfect score: 4008
Sequence: 1 gccgagccgcgcagtagc.....aaaaaaaaaaaaaaaaaaaa 4008

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1152032 seqs, 561743493 residues

Total number of hits satisfying chosen parameters: 2304064

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_New.*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1283	32.0	2082	1	PCT-US03-27411-47
2	248.6	6.2	1914	1	PCT-US03-09929-1
3	248.6	6.2	1914	1	PCT-US03-09929-7
4	248.6	6.2	1914	1	PCT-US03-09929-9
5	247.8	6.2	3112	1	PCT-US03-09929-5
6	246	6.1	2739	1	PCT-US03-09929-11
7	223.2	5.6	2049	1	PCT-US03-09929-3
8	215.2	5.4	2589	1	PCT-US02-36071A-168
9	215.2	5.4	4157	1	PCT-US02-36071A-167
10	205.6	5.1	2586	1	PCT-US02-36071A-165
11	205.6	5.1	2769	1	PCT-US02-36071A-164
12	155.6	3.9	2705	1	PCT-US03-27411-44
13	145	3.6	725	1	PCT-US03-27411-40
14	69.2	1.7	19290	7	US-60-500-337-19393
15	69.2	1.7	46339	7	US-60-500-337-19384
16	66.6	1.7	32874	1	PCT-US02-36071A-154
17	66.2	1.7	96595	1	PCT-US02-38582-43
18	65.2	1.6	349443	5	US-09-947-914-47
19	65	1.6	201	7	US-60-487-610-23416
20	65	1.6	60327	7	US-60-485-450-15208
21	65	1.6	2021	7	PCT-US02-36071A-187
22	64.6	1.6	66843	3	US-60-500-337-19683
23	64.4	1.6	450000	6	US-10-650-449-3
24	64.2	1.6	55492	7	US-60-487-610-19542
25	64.2	1.6	55492	7	US-60-485-450-12034
26	64.2	1.6	56144	7	US-60-495-135-3633

27	64	1.6	67191	1	PCT-US02-38582-169
c 28	64	1.6	8059021	5	US-09-947-914-523
29	63.8	1.6	179338	7	US-60-499-964-5203
30	63.6	1.6	201	7	US-60-487-610-64272
31	63.6	1.6	20690	7	US-60-487-610-19632
32	63.6	1.6	20690	7	US-60-485-450-12090
c 33	63.6	1.6	94381	1	PCT-US02-38582-31
c 34	63.4	1.6	201	7	US-60-499-964-2607
c 35	63.4	1.6	201	7	US-60-499-964-2629
c 36	63.4	1.6	201	7	US-60-499-964-2651
c 37	63.4	1.6	201	7	US-60-499-964-2673
c 38	63.4	1.6	201	7	US-60-499-964-2689
c 39	63.4	1.6	201	7	US-60-499-964-2711
c 40	63.4	1.6	201	7	US-60-499-964-2733
c 41	63.4	1.6	201	7	US-60-499-964-14691
c 42	63.4	1.6	201	7	US-60-499-964-21318
c 43	63.4	1.6	201	7	US-60-500-337-81685
c 44	63.4	1.6	201	7	US-60-500-337-81691
c 45	63.4	1.6	62085	7	US-60-487-610-19240

ALIGNMENTS

RESULT 1
PCT-US03-27411-47
; Sequence 47, Application PC/TUS0327411
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; Lee, Soo Yeun;
; APPLICANT: ELLIOTT, Vicki S.; HAFALIA, April J.A.;
; APPLICANT: GRIFFIN, Jennifer A.; YANG, Yonghong G.;
; APPLICANT: CHAWLA, Narinder K.; BAUGHN, Mariah R.;
; APPLICANT: BECHA, Shanya D.; KHARE, Reena;
; APPLICANT: THORNTON, Michael B.; MASON, Patricia M.;
; APPLICANT: GIETZEN, Kimberly J.; ISON, Craig H.;
; APPLICANT: RAMQUIS, Joseph P.; SWARNAKAR, Anita;
; APPLICANT: RICHARDSON, Thomas W.; TRAN, Uyen K.
; TITLE OF INVENTION: NEUROTRANSMISSION-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-1557 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/27411
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 60/408,383
; PRIOR FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 60/408,781
; PRIOR FILING DATE: 2002-09-05
; PRIOR APPLICATION NUMBER: US 60/414,221
; PRIOR FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/426,483
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 60/431,566
; PRIOR FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: US 60/434,317
; PRIOR FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US 60/437,763
; PRIOR FILING DATE: 2003-01-03
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PERL Program
; SEQ ID NO 47
; LENGTH: 2082
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No: 7523644CB1
PCT-US03-27411-47

Query Match 32.0%; Score 1283; DB 1; Length 2082;
Best Local Similarity 82.5%; Pred. No. 0;
Matches 1546; Conservative 0; Mismatches 215; Indels 114; Gaps 2;
Qy 61 CAGCGGAGCAGAGATGCTTGGCAGGCGCGCCCGCGCGCT 120


```

; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 1
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1908)
PCT-US03-09929-1

```

	Query Match	6.2%	Score 248.6;	DB 1;	Length 1914;
	Best Local Similarity	52.1%;	Pred. No. 1.5e-56;		
	Matches 704;	Conservative 0;	Mismatches 629;	Indels 18;	Gaps 6;
QY	383	GAAGGATCGACTCGATGGTGACCTGACACTCACAGACAGAACTCGAGAGAAAGGCAAG- 441			
DB	200	GAGCGATCTCTGGGAGGCCCGCTGGAGAGAGAGACTGAGTGTATCCAGAAAGGAAGA 259			
QY	442	--AAAGAGGACGAATGTCACAAATTTATCCAGATTTCTCGCCATTTGTCATATGCCCTCTCACC 499			
DB	260	ACAACCAGACCGAGTGTTCACATTTATCCCGCTTCCTTCGACGCCCTACAATGCGCTCCACC 319			
QY	500	TCCTCACTGCGGCACCTTCGCTTTTTCATCCGAAGTCGGGGTATTGATGTGTCCAGATT 559			
DB	320	TGTACGTCTGTGGCACCTTAGCCCTTCAGGCCCAAGTGCACCTACGTCAACATGCTCACCT 379			
QY	560	TCCA---GCAGGTGTAAGACACTTGAGAGCGCCGGGGGAAATGCTCTTTTGAGCCAGCTC 616			
DB	380	TCACTTTGGAGCATGGAGAGTTTGAAGATGGGAAGGCAAGTGTCCCTATGATCCACGCTA 439			
QY	617	AACGGTCAGACGCTGAATGSGCTGGGGCGTCTCTACACCGCCACCTGTGAAAGACTTC 676			
DB	440	AGGGCATGTGTGCCCTTCCTTGTGGATGGTGAGCTGACTCGGCCACACTCAACAACACTTC 499			
QY	677	TGGGGACTGAGCCATCATCTCCCGAGCTGTGGTGCAGCTGAGGACTGGATTCGAACAG 736			
DB	500	TGGGCAGGAACCATATATCTCGGTAAACATGGGCG---CCCACTCATCAAGACAG 556			
QY	737	AGACCTTGTCATCTCGCTTAATGCTCCAGCCTTTGTGCGAGCTATGGTCTTGAGCCAG 796			
DB	557	AGTACCTGGCCTTTTGCTCAACGAACCTCACTTTGTAGGCTCTGCCTATGTACCTGAGA 616			
QY	797	CTCAGTGGGGGATGAAGATGGAGACATGAATATCTTTTTTTCTCACGGAGACCTCC 856			
DB	617	GTGTGGGCAGCTTTCACGGGGGACGACACAAAGTCTACTTCTTCTTCAGGGAGCGGCAG 676			
QY	857	GAGTGTGGACTCCTATGACGGCATCAAGTCCCAAGAGTGGCCCGAGTGTGTGCGGGG 916			
DB	677	TGAGTCCGACTGCTATCGCGCAGAGTGGTGGCTCGTGTGGCCCGTGTCTGCAAGGGCG 736			
QY	917	ACCTTGGGGCAGGAAGACCTTTTCAGCAGAGATGGACAGTTCCTGAAGGCTGACCTGC 976			
DB	737	ATATGGGGGCGCAGGACCCCTCGAGGAAAGTGGACACGTTCTCTGAAGGCCGCGCTGG 796			
QY	977	TGTGCCCAGGGCCGAGCATGGCCGGGCTTCGGGGGTTCTGCAAGGCTATGGCAGAGCTTC 1036			
DB	797	CATGCTCTGCCCGAACTGGCAGCTCTACTTCAACCAAGCTGCAAGGGGATG---CACACC 853			
QY	1037	GGCCTCAGCCTGGAGCGGGGAACCCCATCTTTTATGGGATCTTTTCTCTCCAGTGGGAAG 1096			

Db	854	TGCAGGACACCTCTCTGGCACAACACCACTTCCTTTGGGGTTTTTTCAGACACAGTGGGGT	913
Qy	1097	GAGCTGCATCTCTGCTGTGTGTCCTTCGACACCCCAAGACATCCGGGCGAGTGTCTGAATG	1156
Db	914	ACATGTACCTCTCGGCCATCTGTGAGTACCAGTTTGGAAAGAGATCCACGGGGTGTTCGAGG	973
Qy	1157	GTCCTTTTAGAGACTTAAACATGACTGCAACAGGGGACTGCGCTGTTCATGGACACAGG	1216
Db	974	GCCCTATAAGGAGTACCATGAGGAAGCCCAAGAGTGGGACCGCTACACTGACCT--G	1030
Qy	1217	TGCCAGCCAGCAGCTGGAGGTGCATCGCCCAACAATGAGCTCCAGCAGTTTGGAT	1276
Db	1031	TACCCAGCCCTCGCCCTGGCTGTGTCGTATCACTGCGCATTAACAACTGGCATCGCGGCCACGGCTACACCA	1090
Qy	1277	CCTCACCTCTCCCTGCCAGACCGCTGTCTACCTTTTATCAGAGACCAACCCCTCTCATGGACA	1336
Db	1091	GCTCCCTGGAGCTACCCGACAAATCCTCAACTTCGTCAAGAGCACCCCGCTGATGGAGG	1150
Qy	1337	GGCCCGTGTTCGGGCTGACGGCGGCCCTGTGCTGCTACTACAGATACAGCCTTATCTCA	1396
Db	1151	AGCAGTGGGGCCCTCGGTGGAGCGCGCCCTGTCTGCTGAAGAGGGCACCAACTTCACCC	1210
Qy	1397	GAGTCGTGGGCCACAGGCTGACCAAGCCTCTCAGGGAAGAATATGACGTGCTCTACCTGG	1456
Db	1211	ACCTGGTGGCGACCGGTTACAGAGCTGTATGGAGCCACTATACAGTGTGTTTCATG	1270
Qy	1457	GGACAGAGGATGGACACTCTCACCGGCTGTGCGCATTTGGAGCTCAGCTCAGTCTTTGG	1516
Db	1271	GCACAGGAGCGCTGCTCTCAAGGCTGTGAGCCTTGGGCCCTGGGTTCACTGATG	1330
Qy	1517	AGGATGTGGCTTGT--CCAGAACACAGCCGGTTGAGAGCATGAAATGTPACCAG	1573
Db	1331	AGGAGCTGCAGCTCTTTGACGAGGAGCCCATGAGAAGCCTGGTCTATCTCAGAGCAAGA	1390
Qy	1574	ATTGGCTCCTGGTGGGCTCCCATACTCAGGTGACACAAGTCAACACAGCAACTGTGGCC	1633
Db	1391	AGCTGCTCTTTGGCGGCTCCCGCTCTCAGCTGGTGCAGCTGCCGTGGCCGAGTGCATGA	1450
Qy	1634	GTCFCCAGAGCTGCTCGAGTGTATCTGGCCACAGGACCCCGTGTGGCGCTGGAGCTTCC	1693
Db	1451	AGTATCGCTCCTGTGCAGACTGTCTCTCGCCGGGACCCCTATTTCGCTGGAGCGTCA	1510
Qy	1694	GGCTTGATGCTTGTGTGGCCACCGCCGCGA	1724
Db	1511	ACACCAGCGCTGTGTGGCGCTGGGTGGCCA	1541

RESULT 3

PCT-US03-09929-7

; Sequence 7, Application PC/TUS0309929

; GENERAL INFORMATION:

; APPLICANT: Curagen Corporation, et al.

; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, A

; FILE REFERENCE: 21402-573B-061

; CURRENT APPLICATION NUMBER: PCT/US03/09929

; CURRENT FILING DATE: 2003-04-01

; PRIOR APPLICATION NUMBER: 60/368,996

; PRIOR FILING DATE: 2002-04-01

; PRIOR APPLICATION NUMBER: 60/369,980

; PRIOR FILING DATE: 2002-04-04

; PRIOR APPLICATION NUMBER: 60/370,381

; PRIOR FILING DATE: 2002-04-05

; PRIOR APPLICATION NUMBER: 60/370,969

; PRIOR FILING DATE: 2002-04-08

; PRIOR APPLICATION NUMBER: 60/371,002

; PRIOR FILING DATE: 2002-04-09

; PRIOR APPLICATION NUMBER: 60/372,002

; PRIOR FILING DATE: 2002-04-12

; PRIOR APPLICATION NUMBER: 60/384,297

; PRIOR FILING DATE: 2002-05-30

; PRIOR APPLICATION NUMBER: 60/386,816

; PRIOR FILING DATE: 2002-06-07

; PRIOR APPLICATION NUMBER: 60/389,123

```
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 7
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1914)
PCT-US03-09929-7

Query Match          6.2%; Score 248.6; DB 1; Length 1914;
Best Local Similarity 52.1%; Pred. No. 1.5e-56;
Matches 704; Conservative 0; Mismatches 629; Indels 18; Gaps 6;

QY 383 GAAGATCGACTGGATGTTACCTGAGACTCAGACAGAACTGCAGGAAGAAAGCAAG- 441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 200 GAGCGATCTCTGGGAGGCCCGGTGGAGGAGAACTGAGTGTATCCAGAAAGGAAGA 259

QY 442 --AAGAGACGAATGTCAAAATTTTATCCAGATTTCTCGCCATGTGTAATGCTCTACC 499
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 260 ACAACCCAGACCGAGTGTCTCACTTCACTCCGCTTCTCGAGCCCTACATGCTCCACC 319

QY 500 TCCTCAGCTGGCGGACCTTTCGCTTTTGATCCGAGTGGGGTTATTGATGTGTCCAGTT 559
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 320 TGTACGTGTGGGCACTTACGCCCTTCAGCCCAAGTGCACCTACGTCAACATGCTCACT 379

QY 560 TCCA---CGAGTTGAAAGACTTTGAGAGCGCGCGGGGAAATGCTCTTTTGGAGCAGTC 616
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 380 TCACTTTGGAGATGGAGATTTGAAGATGGAGGGCAAGTGTCCCTATGACCCAGCTA 439

QY 617 AACGTCAGCAGCTGTAATGCTGGGGCGTCTCTACACCGCACTGTGAAGAACTTCC 676
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 440 AGGGCCATGCTGGCCCTTCTTGGATGGTGAGCTGTACTCGGCCACACTCAACAACCTCC 499

QY 677 TGGGACTGAGCCCATCATCTCCCGAGCTGTGGTGGAGCTGAGGACTGGATTCCGAACAG 736
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 500 TGGGACGGACCCCATTTCTGCTAACATGGGGC---CCACCACTCCATGAAGACAG 556

QY 737 AGACCTTGTCATCTCGGTTTAATGTCAGGCTTTTGTGCGAGCTATGTCTCTGAGCCGAG 796
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 557 AGTACTCGGCTTTTGGCTCAACGAACCTCACTTTGTAGGCTCTGCTATGTACCTGAGA 616

QY 797 CTGAGTGGGGGATGAAGATGGAGACGATGAATCTTTTCTTCTACGGAGACCTCCC 856
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 617 GTGTGGGAGCTTACGGGGGACGACGACAGAGTCTACTTCTTTCAGGGAGCGGGCAG 676

QY 857 GAGTGTGGGACTCTCTATGAGGCGCATCAAGGTCCCAAGAGTGGCCCGAGTGTGTGGGGGG 916
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 677 TGGAGTCCGACTCTATCCGAGCAGGTGGTGGCTGTGTGGCCGCTGTCTGCAAGGGG 736

QY 917 ACCTTTGGGGCAGGAAGACCTTTCAGCAGAGATGAGCAGCTTTCTGAAGGCTGACCTGC 976
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 737 ATATGGGGGGCGACGGACCTTCAGAGAGAGTGGACCACTCTCTCTGAAGGCGCGCTGG 796

QY 977 TGTGCCAGGGCCCGAGCATGGCCGGGCTCCGGGTTCTGCAGGCTATGGCAGAGCTTC 1036
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 797 CATGCTGTGCCCGCACTGGCAGCTCTACTTCAACAGCAGCTGCAGCGGATG---CACACC 853

QY 1037 GGCCTCAGCCTGGAGCGGAACCCCATCTTTTATGGATCTTTTCTCCAGTGGGAAG 1096
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 854 TGCAGGACACCTCTCTGGCAGACACACCTCTTTTGGGGTTTTCACGACAGTGGGGTG 913

QY 1097 GAGCTGCCATCTCTCTGTGTGCTTCCGACCCCAAGACATCCGGGAGTGTGTAAG 1156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 914 ACATGTACTGTGCGGCCATCTGTGAGTACCACTTGTGAAGAGATCCAGCGGGTGTGAGG 973

QY 1157 GTCCTTTAGAGACTAATACATGACTCAACAGGGGACTGCCGTGTCATGACACAGGAG 1216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 4

PCT-US03-09929-9

; Sequence 9, Application PC/TUS0309929

; GENERAL INFORMATION:

; APPLICANT: Curagen Corporation, et al.

; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND MET

; FILE REFERENCE: 21402-5738-061

; CURRENT APPLICATION NUMBER: PCT/US03/09929

; CURRENT FILING DATE: 2003-04-01

; PRIOR APPLICATION NUMBER: 60/368,996

; PRIOR FILING DATE: 2002-04-01

; PRIOR APPLICATION NUMBER: 60/369,980

; PRIOR FILING DATE: 2002-04-04

; PRIOR APPLICATION NUMBER: 60/370,381

; PRIOR FILING DATE: 2002-04-05

; PRIOR APPLICATION NUMBER: 60/370,969

; PRIOR FILING DATE: 2002-04-08

; PRIOR APPLICATION NUMBER: 60/371,002

; PRIOR FILING DATE: 2002-04-09

; PRIOR APPLICATION NUMBER: 60/372,002

; PRIOR FILING DATE: 2002-04-12

; PRIOR APPLICATION NUMBER: 60/384,297

; PRIOR FILING DATE: 2002-05-30

; PRIOR APPLICATION NUMBER: 60/386,816

; PRIOR FILING DATE: 2002-06-07

; PRIOR APPLICATION NUMBER: 60/389,123

; PRIOR FILING DATE: 2002-06-13

; PRIOR APPLICATION NUMBER: 60/402,207

; PRIOR FILING DATE: 2002-08-09

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 179

; SOFTWARE: CuraseqList version 0.1

; SEQ ID NO 9

; LENGTH: 1914

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1914)
PCT-US03-09929-9

Query Match 6.2%; Score 248.6; DB 1; Length 1914;
Best Local Similarity 52.1%; Pred. No. 1.5e-56;
Matches 704; Conservative 0; Mismatches 62; Indels 18; Gaps 6;

QY 383 GAAGGATCGACTGGATGCTAGCTGAGACTCAGACAGCAAACTGCAGGAGAAAGGCAAGA 442
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 200 GAGCGATCTCTGGGAGGCCCGCCGTTGAGAGAGAGACTGAGTGTATCCAGAAAGGGAAGA 259
443 A---AGAGGAGCAATGTACAAATTTATCCAGATTCTGCCATTTGCTCAATGCTCTCACC 499
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 260 ACAGCCAGCAGAGTGCTTCAACTTCATCCGCTCTCTGCAGCCCTACAAATGCTCTCCACC 319
500 TCCTCAGCTGGGCACTTCGCTTTTGATCCGAAGTGGGGGTTATTGATGTGCCAGTT 559
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 320 TGTACGCTGTGGCACCCTACGCTTCCAGGCCAAGTGCACTACGTAACATGCTCACCT 379
560 TCCA---CCAGGTTGAAGACTTGAGACCGCGCGGGGAAATGTCCTTTTGAGCCAGCTC 616
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 380 TCACCTTTGGAGCATGGAGAGTTTGAAGATGGGAAGGCAAGTGTCCCTATGACCCAGCTA 439
617 AACGTCACGAGCTGTAATGCTGGGGCGTCTCTACACGCCCACTCTGAAGACTTCC 676
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 440 AGGGCATTGCTGGGCTCTTGTGATGTGAGCTGTACTCGGCCACACTCAACAATCTCC 499
677 TGGGACTGAGCCCATCATCTCCGAGCTGTGGTGCAGCTGAGACTGGATTTCGAACAG 736
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 500 TGGGCACGAACCCATTATCTGCGTAACATGGGCG---CCACACACTCCATGAAGACAG 556
737 AGACCTTGTATCTCTGCTTAATGCTCCAGCTTTGTGCGAGCTATGTCCTGAGCCAG 796
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 557 AGTACCTGGCCTTTTGGTCAACGAACCTCTTTGTAGGCTCTGCCCTATGTACCTGAGA 616
797 CTGAGTGGGGGATGAAGATGAGACGATGAATCTTTTCTTCACGGAGACTCC 856
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 617 GTGTGGGAGCTTACGGGGACGACGACAGGTCTACTTCTTTCAGGAGCGGGGAG 676
857 GAGTGTGGACTCTCTATGAGCGCATCAAGGTCCCAAGAGTGCCCGAGTGTGTGGGGGG 916
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 677 TGGAGTCGAGCTGCTATCCGAGCAGGTGGTGGCTCGTGTGGCCGCTGCTGCAAGGCG 736
917 ACCTTGGGGGAGGAGACCTTTCAGCAGAGATGGACGACGTTTCTGAAGCTGACCTGC 976
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 737 ATATGGGGGGCGCAGGGACCTTCAGAGGAAGTGGACCACTTCTGAAGCGCGGCTGG 796
977 TGTGCCAGGCGCCGAGCATGGCCGGGCTCCGGGTTCTGCAGGCTATGCGACGCTTC 1036
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 797 CATGCTCTGCCGAGCTGGCAGCTCTACTTCAACCAGCTCAGGCGGATG---CACACC 853
1037 GGCTTCAGCTGGAGCGGGAACCCCATCTTTTATGGGATCTTTTCTCCAGAGTGGGAAG 1096
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 854 TGCAGGACACCTCTCTGGCACAACACCACTTCTTTGGGGTTTTCAGCAGAGTGGGTG 913
1097 GAGTGGCCATCTCTGCTGTGTGCTTCCGACCCCAAGATCCGGGCGAGTGTGAATG 1156
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 914 ACATGTACTGTGGGCCATCTGTGAGTACCAGTTTGAAGAGATCAGCGGGGTGTTGAGG 973
1157 GTCCTTTTAGAGATAAAGATGACTGCAACAGGGGACTGCCCTGCTATGGACACAGG 1216
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 974 GCCCTATAGAGTACCATGAGGAAGCCAGAGTGGGACCGCTACACTGACCT---G 1030
1217 TGCCCCAGCCAGCTGGAGAGTGCATCGCCACAACATGAAGCTCCAGCAGTTTGGAT 1276
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1031 TACCCAGCCCTCGGCCCTGCTGCTGCTGATTAACAACTGGCATCGGGCCACGCTACACCA 1090
1277 CCTCACTCTCCCTGCCAGCGCTGCTCAGCTTTATCAGAGACACCCCTCTCATGGACA 1336
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 1091 GCTCCCTGGAGCTACCCGACAACATCCTCAACTTCGTCAAGAAGCACCCGCTGATGGAGG 1150
QY 1337 GGCCGCTGTCCCGGCTGAGCGGCCCGCCCTGCTGCTCACTACATACAGCTATATCTCA 1396
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1151 AGCAGGTGGGCGCTCGGTGGAGCCCGCCCTGCTGTAAGAGAGGCGCAACACTTCACCC 1210
1397 GAGTGTGGCCACACAGGTGACCAGCTCTCTCAGGAGAAATATACAGTGTCTTACCTGG 1456
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1211 ACCTGGTGGCGACCGGGTTACAGGACTTGTATGGAGCCACCTATACAGTGTCTTCATTG 1270
1457 GGACAGAGATGGACACCTCCACCGGGCTGTCGGCATTTGGAGCTCAGCTCAGTGTCTGG 1516
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1271 GCACAGAGAGCGGTGGCTGCTCAAGGCTGTGAGCTGGGGCCCTGGGTTCACCTGATTG 1330
1517 AGGATCTCGCCTTGT---CCAGAACACACAGCCGCTTGAGAGCATGAAATGTATCCAGG 1573
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1331 AGGAGCTCAGCTGTTTGCACAGAGCCCATGAGAAGCCTGGTGTATCTCAGAGCAAGA 1390
1574 ATTGCTCCTGTGGGCTCCCATACTGAGGTGACAAAGTGAACACCACTGAGCTGTGGCC 1633
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1391 AGCTGCTCTTTGCCGCTCCGCTCTCAGCTGCTGAGTGTGCGCCGCTGCGCCGACTGCATGA 1450
1634 GTCTCCAGAGCTGCTCGGAGTGTATCTTGGCCAGGACCCCGTGTGGCCCTGGAGCTTCC 1693
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1451 AGTATCGTCTGTGCAGACTGTGTCTCGCCCGGGACCCCTATTTCGCCCTGGAGCGTCA 1510
1694 GGCTTGATGCTTGTGTGGCCACGCGCGGA 1724
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1511 ACACGAGCGCTGTGTGGCCGCTGGGTGGCCA 1541

RESULT 5

PCT-US03-09929-5
; Sequence 5, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND MET
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 5
; LENGTH: 3112
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (104)..(2602)
PCT-US03-09929-5

Query Match 6.2%; Score 247.8; DB 1; Length 3112;

Best Local Similarity 52.1%; Pred. No. 3.5e-56;		Matches 702; Conservative 0; Mismatches 627; Indels 18; Gaps 6;	
QY	387	GATCGACTGGATGGTACTGAGACTCACAGACAGAACTGCAGGAAGAAGGCAAG---AA	443
Db	361	GATCTCTGGGAGGCCCGCTGGAGAGAGAGACTGAGTGTATCAGAAAGGGAACAA	420
QY	444	AGAGCGCAATGTCCACAATTTATTCAGATTCTGCCAATTGTCAATGCCCTCTCACCTCCT	503
Db	421	CCAGACCGAGTGGCTTCAACTTCATCCGCTCTCTGCAGCCCTACAATGCCCTCCACCTGTA	480
QY	504	CACGTGGGACACCTTCGCTTTTGTATCGAAGTGGGGGTTATTGATGTGCCAGTTTCCA	563
Db	481	CGTCTGTGGCACCTTACGCCCTTCAGGCCCAAGTGCACCTAGTCAACATGCTCACCTTTCAC	540
QY	564	---CGAGGTTGAAGAGCTTGAGAGCGCCGGGGAAATGTCTTTTCAGACCACTCAACG	620
Db	541	TTTGAGCAATGGAGAGTTTGAAGATGGGAAGGCAAGTGTCCCTATGACCCAGCTAAGG	600
QY	621	GTACGAGCTGTAAATGGCTGGGGCGTCTCTATACAGCCCACTGTGAAGAACTTCTCTGGG	680
Db	601	CCATGCTGGCCCTCTTGTGGATGGTAGCTGTACTCGGCCACACTCAACAACCTTCCTGGG	660
QY	681	GACTGAGCCCATCATCTCCGAGCTGTGGSTCGAGCTGAGAGCTGAGATTGCAACAGACAG	740
Db	661	CACGGAACCATATATCTGGGTAACTATGGGCG---CCACCACTCCATGAAGACAGAGTA	717
QY	741	CTTGTCACTCGGCTTAATGCTCCAGCGTTTGTGCGAGCTATGGTCTCTGAGCCACGCTGA	800
Db	718	CCTGGCCCTTTGGCTCAAGCAACCTCACTTTGTAGGCTCTGCCCTATGTACTGAGAGTGT	777
QY	801	GTGGGGGATGAAGATGGAGACGATGAAATCTTTTTTTTCTCACGAGAGACTCCCGAGT	860
Db	778	GGGCAGCTTCAGGGGGACGACGACAGGTCTACTTCTTCTCAGGAGCGGCGAGTGA	837
QY	861	GTTGGACTCTATGAGCGCATCAAGTCCCAAGATGGCCCGAGTGTGTGCGGGGACCT	920
Db	838	GTCCGACTGTATGCCGAGCAGGTGTGGTCTGTGTGGCCCGTGTCTGCAAGGGCGGATAT	897
QY	921	TGGGGCGAGGAAGACCTTCAGCAGAGATGGACGACGTTTCTGAAGGCTGACCTGCTGTG	980
Db	898	GGGGGGCGACGACCTTCAGAGGAAGTGGACCACTTCTTCTGAGGCGGGCTGGCATG	957
QY	981	CCCAGGCCGAGCATGCGCCGGCCCTCCGGGTTCTCAGGGCTATGCGAGCTTCGGCC	1040
Db	958	CTCTGCCCCGAACCTGGCAGCTCTACTTCAACACAGCTCGAGCGCATG---CACACCTGCA	1014
QY	1041	TCAGCTTGGAGGGGAACCCCATCTTTTATGGATCTTTTCTCCAGCTGGGAGGAGC	1100
Db	1015	GGACACCTCTGGCACAACACCACTCTTTTGGGGTTTTTCAAGCAGAGTGGGGTGACAT	1074
QY	1101	TGCCATCTCTGCTGTGTGCTTCCGACCCCAAGACATCGGCGAGTGTCTGAATGGTCC	1160
Db	1075	GTACTCTGGCCATCTGTGAGTACCAGTTGGAGAGATCCAGGGGTGTTTGGGGCCC	1134
QY	1161	CTTTPAGAGACTAAACATGACTGCAACAGGGAGTGCCTGTATGCAACAGAGTGGC	1220
Db	1135	CTATAAGGAGTACCATGAGGAAGCCCAAGAGTGGGACCCGCTACACTGAACCT---GTACC	1191
QY	1221	CCAGCCAGACTGGAGAGTGCATCGGCCACACATGAAGCTCGACAGATTGGATGCTC	1280
Db	1192	CAGCCCTCGGCCCTGGCTGTGTCATTAACAACCTGGCATCGCGCCAGCGCTTACACCACTC	1251
QY	1281	ACTCTCCCTGCCAGACCGCTGTCACTTTATCAGAGACCACTCTCATGACAGGCC	1340
Db	1252	CTGGAGACTACCCGACAACATCTCAACTTCGTCAAGAAGCAACCGCTGTATGAGGAGCA	1311
QY	1341	CGTGTTCGCCGCTGACGGCCGCCCTGCTGGTCACTTACAGATACAGCTTATCTCAGAGT	1400
Db	1312	GGTGGGCCCTCGGTGGAGCCGCCCTGCTGCTGTGAAGAAGGCAACCACTTCAACCACT	1371
QY	1401	CGTGGGCCACAGGGTGACAGCCTCTCAGGGAAGAATATGAGTGTCTTACCTGGGGAC	1460

```

Db 1372 GGTGGCCGACCGGGTTACAGACTTGTGAGGACCCACCTATACAGTGTGTTTCATTTGGCAC 1431
QY 1461 AGAGATTGGACACCTCCACCGGGCTGCGCATTTGGAGCTCAGTCTCAGTGTCTTGGAGGA 1520
Db 1432 AGGAGACGGTGGCTGCTCAAGGCTGTGAGCCTGGGCCCTGGGTTTCACTGATTGAAGA 1491
QY 1521 TCTGGCCCTTGT---CCAGAACACACGCGGTTTGAGAGCATGAAATTTGTACACGATTG 1577
Db 1492 GCTCAGCTGTTTACCCAGGAGCCCATGAGAAGCCTGGTGTATCTCAGAGCAAGAAGCT 1551
QY 1578 GCTCCTGGTGGGCTCCATCTAGTGTGACACAAGTGAACACACGCAACTGTGGCCGCTCT 1637
Db 1552 GCTCTTTGGCGGCTCCGCGCTCTCAGCTGGTGCAGCTGCCGTGGCCGCGCATGAAGTA 1611
QY 1638 CCAGAGCTGCTCGGAGTGTATCTCTGGCCGACGAGACCCGTGTGCGCCTGGAGCTTCCGCGCT 1697
Db 1612 TCGTCTCTGTGCAGACTGTCTCTCGCCCGGAGCCCTATTTGCCCTTGGAGCGTCAACAC 1671
QY 1698 TGATGCTTGTGTGGCCCAACGCGCGCA 1724
Db 1672 CAGCCGCTGTGTGCGCGTGGGTGGCCA 1698

RESULT 6
PCT-US03-09929-11
; Sequence 11, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND MET
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: Curaseq1 version 0.1
; SEQ ID NO 11
; LENGTH: 2739
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (104)..(2602)
PCT-US03-09929-11

Query Match 6.1%; Score 246; DB 1; Length 2739;
Best Local Similarity 51.8%; Pred. No. 9.9e-56;
Matches 708; Conservative 0; Mismatches 640; Indels 18; Gaps 6;

QY 368 CTGGGGAAGACCCCGAAGATCAGCTGGATGTCTACCTGAGACTCAGACAGCACTGCA 427
Db 342 CTCCTGGAGTGCAGGAGGACGATCTCTGGAGGCCCGCTGGGAGAAGAAGACTGAGTGA 401
QY 428 GGAAGAAGCAAG---AAAGAGGACGAATGTCCACAAATTTTATCCAGATTTCTGCCAATTG 484

```


Db 440 TCACCTTTGAGCATGGAGAGTTGAAGATGGGAAGGCAAGTGTCCCTATGACCCAGCTA 499
QY 617 AACGGTCAGACAGCTGTAATGGCTGGGGGCTCTCTACACCGCCAGCTGTGAAGAAGTCC 676
Db 500 AGGGCCATGCTGGCCTTCTGTGGATGGTGAAGTGTACTGCGCCACACTCAACAAGTCC 559
QY 677 TGGGGACTGAGCCCATATCTCCCGAGCTGTGGGTGAGCTGAGGACTGGATTCGAACAG 736
Db 560 TGGGACGGAACCCATTATCTCGCTGAACATGGGCG--CCCAACCACTCATGAAGACAG 616
QY 737 AGACCTTGCATCTCGGCTTAATGCTCCAGCCTTTGTCCAGCATATGGTCTCGAGCCAG 796
Db 617 AGTACCTGCCCTTTGGCTACAGAACCTCACTTTGTAGGCTCTGCCTATGACTCTGAGA 676
QY 797 CTCAGTGGGGGGTGAAGATGGAGACGATGAATCTTTTTTCTTACCGGAGACCTCC 856
Db 677 GTGTGGGAGCTTCACGGGGGAGCAGACAAAGTCTACTTCTTTCAGGGAGCGGAG 736
QY 857 GAGTGTGAGCTCTATGAGCGCATCAAGGTCCCAAGAGTGGCCGAGTGTGTGGGGGG 916
Db 737 TGGAGTCCACACTGCTATGCGAGCAGGTGGTGGCTGCTGTGGCCGCTGTCTGCAAGGGG 796
QY 917 ACCTTGGGGGAGGAGACCTTTCACAGAGATGACGACGTTTCTGAAAGGCTGACCTGC 976
Db 797 ATATGGGGGCGCAGGACCTCGAGAGGAAGTGGACCACTTCTTGAAGGCGCGGCTGG 856
QY 977 TGTGCCAGGGGCGCAGCATGCGCGGGCTCTCGGGGTTCTCGAGGCTATGCGAGAGCTTC 1036
Db 857 CATGCTCTGCCCGAAGTGGCAGCTACTTCAACCAAGCTGCAGGGATG---CACACC 913
QY 1037 GGCCTAGCCTGGAGGGGAAACCCCATCTTTATGGGATCTTTTCTCCCAAGTGGGAAG 1096
Db 914 TGCAGGACACCTCTCGGCAACACACACCTTCTTTGGGGTTTTCAGGACACAGTGGGGT 973
QY 1097 GAGCTGCCATCTGCTGTGTGCTTCCGACCCCAAGACATCCGGGAGTGTCTGAATG 1156
Db 974 ACATGTACCTGTGGCCCATCTGTGAGTACCACTTGAAGAGATCCAGCGGGTGTGGAG 1033
QY 1157 GTCCCTTTAGAGAGCTTAAACATGACTGCAACAGGGGAGTGCCTGTCTATGGACACGAGG 1216
Db 1034 GCCCTATAAGAGTACCATGAGGAAGCCAGAGTGGGACCGCTACACTGACCT---G 1090
QY 1217 TGCAGGAGGAGCTGAGAGTGCATGCGCAACAAATGAAGTCCAGCAGTGTGGAT 1276
Db 1091 TACCCAGCCCTCGGCTGCTGCTGCTGATTAACAAGTGGCATCGGGGCCACGGCTACACCA 1150
QY 1277 CTTCACTCTCCCTGCGAGACCGGCTGCTCACCTTTATCAGAGACCACTCTCATGGACA 1336
Db 1151 GTCCTGTGAGTACCCGAAACATCTCTCACTTCTGTAAGAGCACCCGCTGATGGAGG 1210
QY 1337 GGCCTGTGTCCGCTGACGGCCGCCCCCTCTGCTGCTCACTACAGATACAGCCTATCTCA 1396
Db 1211 AGCAGTGGGGCTCGGTGGAGCCGCCCTCTCTGTAAGAGGCGCACTTCACTCC 1270
QY 1397 GAGTGTGCGCCACAGGTGACAGCCTCTCAGGGAAGAATATGAGTGTCTACCTGG 1456
Db 1271 ACCTGTGGCCGACCGGCTTACAGGACTGTATGGAGCCACCTATACAGTGTGTTCATG 1330
QY 1457 GCACAGAGGATGACACCTCCACCGGCTGTGCGCATTTGGAGCTCAGCTCAGTGTGG 1516
Db 1331 GCACAGGAGCGGCTGGAGCTCAAGGCTGTGAGCCTGGGGCCCTTCACTGATG 1390
QY 1517 AGGATCTGGCCCTGTT 1532
Db 1391 AGGAGCTGACGCTGTT 1406

RESULT 8

PCT-US02-36071A-168

; Sequence 168, Application PC/TUS0236071A

; GENERAL INFORMATION:

; APPLICANT: Sagres Discovery

; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000142
; CURRENT APPLICATION NUMBER: PCT/US02/36071A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: US 10/052,482
; PRIOR FILING DATE: 2001-11-08
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 168
; LENGTH: 2589
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US02-36071A-168

Query Match 5.4%; Score 215.2; DB 1; Length 2589;
Best Local Similarity 48.9%; Pred.No. 2.3e-47;
Matches 747; Conservative 0; Mismatches 748; Indels 33; Gaps 5;

QY 211 CCAGAACCTCGCTGCCCATCTCCGAGGCTGACTCTATCTCACCCGGTTTGCAGCGTCT 270
Db 76 CCATATACCCCGGATCACCTGGGAGCACAGAGGTGCACCTGGTGAGTTTCATGAGCCA 135
QY 271 CATAGCTACAATTAATCT 330
Db 136 GACATCTACAACCTCTAGCCTTGTCTGTGAGCGAGGACAAAGACACCTTGTACATAGT 195
QY 331 GCACGGGATAGCATCTTCGCTTTAAACCTCCCTCTCTCTCTCTCTCTCTCTCTCTCT 390
Db 196 GCGCGGGAGGGGCT 255
QY 391 GACTGGATGGTACCTGAGACTCACAGACAGAACTGCAGGAAGAAAGGCAAG---AAAGAG 447
Db 256 TATTGGAAGTCTCAGAGACAAAAGCAAAATGTGCAGAAAGGGAATCAAAACAG 315
QY 448 GACGAATGTCAAAATTTATCCAGATTCGCGCATTTGCAATGCTCTCACCTCTCTCAG 507
Db 316 ACAGAGTGCTCAAACTACATCGGGTGTCTGACGCCACTCAGGCGCCACTTCCCTTTACGT 375
QY 508 TCGGCACCTTCGCTTTTGTATCCGAAGTCGGGGTTATTGATGTGTCAGTTTCCAGCAG 567
Db 376 TGTGGACCAAGCATTCACGCGGCTGTGACCACTTAACTTAACTCTTTAAGTTT 435
QY 568 GTTGAAGACTTGAAGCGCGCGGGGAATTCCTTTTGAAGCCAGCTCAACGGTCAGCA 627
Db 436 CTGGGAAAAATGAAGATGGCAAGAGATGTCCCTTTGACCCAGCACAGCTACACA 495
QY 628 GCTGTAAAGTGGGGGCTCTCTACACCGCCACTGTGAAGAACTTCTCTGGGAGTGA 587
Db 496 TCCGTATGTTGATGGAGAACTTTATTCGGGAGCTCGTATATAATTTTGGGAAGTGA 555
QY 688 CCATCATCTCCGAGCTGTGGTGCAGCTGAGGACTGGATTCGAACAGAGACCTTGTCA 747
Db 556 CCATCATCTCCGAAATTCCTCCACAGCTCTCTGAGGACAGATATCAATTC--- 610
QY 748 TCCTGGCTTAATGCTCCAGCCTTTGTGCGAGCATATGTCCTGAGCCAGCTGAGTGGGG 807
Db 611 -CTTGGCTGAACGAGCTAGTTTCGTTGTTGCTGAGCTGATCCGAAAAAGCCAGACAG 669
QY 808 GATGAAGTGGAGAGATGAATCTTTTTCCTACCGGACCTCCCGAGTGTGGAG 867
Db 670 CCCGCGGAGATGACAGGCTTACTTCTTCTCACGAGGTGCTGTGTGAGTATGAG 729
QY 868 TCCTATGAGCGCATCAAGTTCCAAGAGTGGCCGAGTGTGTGCGGGGACCTTGGGGC 927
Db 730 TTTGTTCAGGGTCTGATCCAGGATAGCAAGAGTGTCAAGGGGACCGAGCGGC 789
QY 928 AGGAAGACCTTTCAGCAGAGATGAGAGCTTTCTTAAGGCTGACCTGCTGCCAGGG 987
Db 790 CTGAGGACCTTGCAGAAGAAATGAGACCTCTCTTCTTCTTCTTCTTCTTCTTCTTCT 849
QY 988 CCCGAGATGCGCGGCGCTCCGGGGTCTTGTGAGGCTATGGCAGAGCTTCGGGCTCAGCT 1047

Db 850 CCAGACAGCGGCTGGTCTTCAATGTGCTCGGGAATGCTTCTGCTCAGGTC---CCCG 906
Qy 1048 GGAGCGGGAACCCCACTCTTTTATGGATCTTTTCTCCAGTGGGAAGAGTGCATC 1107
Db 907 GGCTGAAGTGCGCTGTGTCTATGCACTCTTACCCCAAGTGAACAAGTGGGGCTG 966
Qy 1108 TCTGCTGTGTGTGCC-----TTCCGACCCCAAGACATCCGGGCGAGTGCCTGAATGGT 1158
Db 967 TCGCAGTGTGCGGCTACACCTGTCCACAGCGAGGAGTCTTCTCCACGGGAAGTAC 1026
Qy 1159 CCCTTTAGAGAGTAAACATGACTGCAACAGGGAGTGCCTGTCTATGGACAACAGAGTG 1218
Db 1027 ATGACAGACACACAGTGGAGCAGTCCACACCAAGTGGTGGCTATATATGGCCGGTA 1086
Qy 1219 CCCAGCCAGACTGGAGAGTGCATCGCCCAACAAATGAAGCTCCAGCAGTTTGGATCC 1278
Db 1087 CCAAGCGGGGCTGGAGGTGCATCGACAGAGGAGCGCGGCGCAACTACACAGC 1146
Qy 1279 TCACTCTCCCTGCGACAGCGCTGCTCACCTTTATCAGAGACCAACCTCTCATGGACAGG 1338
Db 1147 TCCTTGAATTTGCCAGACAGACGCTGCAGTTCGTTAAAGACCAACCTTTGATGGATGAC 1206
Qy 1339 CCGGTGTTCCGGCTGAGGCGCGCCCTGCTGCTGCTACATACAGCCTATCTCAGA 1398
Db 1207 TCGGTAAACCCCAATAGACAAACAGGCCCAAGGTTAATCAAGAAAGATGTGAACACTACACCAG 1266
Qy 1399 GTGTCGCCACAGAGGTGACACGCTCTCAGGGAAGAAATATGACGTGCTTACCTGGG 1458
Db 1267 ATCGTGTGACCGGACCCAGGCCCTGGATGGGACTGCTATGATGTCATGTTGTGACG 1326
Qy 1459 ACAGAGATGGACACCTCCACCGGGCTGCGCAATTGGAGCTCAGCTCAGTGTCTTGGAG 1518
Db 1327 ACAGCGGGGAGCTCTGCACAAAGCCATCAGCCTCGAGCACGCTGTTACATCATCGAG 1386
Qy 1519 GATCTGGCCTTTCCAGAACACAGCCGCGTGTAGAGCATGAATTTGTA----- 1568
Db 1387 GAGACCCAGCTCTTCCAGGACTTTGAGCCAGTCCAGACCCCTGCTGTCTTCAAGAAG 1446
Qy 1569 --CCAGATTGGCTCTGTGGGCTCCCATACTGAGGTGACAAAGTGAACACAGCAAC 1626
Db 1447 GGCAACAGTTTGTCTATGCTGCTTAACCTCGGGGTGTCAGGCCCGCTGGCCTTC 1506
Qy 1627 TGTGGCGCTCCAGAGCTCTCGAGTGTATCTGSCCAGGACCCCGTGTGCGCCTGG 1686
Db 1507 TGTGGGAAGCACGCACTTCGAGGACTGTGTGTCGCGGGGACCCCTACTGCGCCTGG 1566
Qy 1687 AGCTTCGGCTTGTATGCTTGTGGGCC 1714
Db 1567 AGCCCGCCACAGCGACCTCGGTGGCTC 1594

RESULT 9

PCT-US02-36071A-167
; Sequence 167, Application PC/TUS0236071A
; GENERAL INFORMATION:
; APPLICANT: Sagres Discovery
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000142
; CURRENT APPLICATION NUMBER: PCT/US02/36071A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: US 10/052,482
; PRIOR FILING DATE: 2001-11-08
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 167
; LENGTH: 4157
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US02-36071A-167

Query Match 5.4%; Score 215.2; DB 1; Length 4157;
Best Local Similarity 48.9%; Pred. No. 3.3e-47;
Matches 747; Conservative 0; Mismatches 748; Indels 33; Gaps 5;
Qy 211 CCAGAACCTCGCTGCCCATCTCGAGGCTGACTCTCTATCTCACCOCGGTTTGACGCTCT 270
Db 163 CCATACCCCGGATACCTCGGAGCACAGAGAGTGCACCTGGTGTGATTTTCATGACCCA 222
Qy 271 CATACGTACAAATATCTGCTCTCTTGTGATCCTCCCTCCACACACTTTACGTCGGT 330
Db 223 GACATCTACAACTACTCAGCCTTGTCTGCTGAGGAGGAGAACACCTTGTATAGGT 282
Qy 331 GCACGGATAGCATCTTCGCTTTAAACCTCCCTCTCTTGGGAAAGACCCCGAAGATC 390
Db 283 GCCGGGAGCGGCTCTCGCTGTGAAGCACTCAACATCTCCGAGAGCAGCATGAGGTG 342
Qy 391 GACTGGATGTACTGTAGACTCAGACAGAACTGCAGGAGAGAAAGCAAG---AAAGAG 447
Db 343 TATTGGAAGTCTCAGAAGACAAAAGCAAAATGTSCAGAAAGGGGAAATCAAAACAG 402
Qy 448 GACGAATGTACAAATTTATCCAGATTCTCGCATTTCTCAATGCTCTCACTCTCTCACG 507
Db 403 ACAGATGCTCAACTACATCCGGGTCTCAGCCACTCAGGCCACTTCCCTTACGTG 462
Qy 508 TCGCGCACCTTCGCTTTTGTATCCGAAGTGCAGGTTATTTGATGTGTCCAGTTTCCAGCAG 567
Db 463 TGTGGGACCAACGATTCCAGCGGCTGTGACCACCTGAACCTTAACATCTCTTTAACTTT 522
Qy 568 GTTGAAGACTTGAGAGCGCGCGGGAATGTCTTTTGGAGCCAGCTCAACGTTGACCA 627
Db 523 CTGGGGAATAATGAAGATGGCAAGGAAGATGTCTTTTACCCAGCACACAGTACACA 582
Qy 628 GCTGTATGTGCTGGGCGCTCTTACACCGCACTGTGAAGAACTTCTCTGGGAGCTGAG 687
Db 583 TCGTCTATGTTGATGAGAACTTTATTCGGGAGCTGCTATTAATTTTGGGAAGTGA 642
Qy 688 CCATCATCTCCGAGCTGTGGTTCGAGCTGAGACTGGAATCGAAGACAGACCTTGTCA 747
Db 643 CCATCATCTCCGAAATTTCTCCACAGTCTCTGAGGACAGAAATGAATCTC----- 697
Qy 748 TCTGCTTAATGCTCCAGCTTTGTGCGAGCTATGCTCTGAGCCAGCTGAGTGGGG 807
Db 698 -CTGGTGAACGAGCCTAGTTTCTGCTGACGTGATCGAAGAGCCAGCTCAGTCCAGCAG 756
Qy 808 GATGAAGATGGAGACGATGAATCTTTTCTTCTACGAGACCTCCCGAGTCTTGGAC 867
Db 757 CCGACGCGGAGATGACAGGCTTACTTCTTTCAGGAGGTGTCTGTGGATATGAG 816
Qy 868 TCCTATGAGCGCATCAAGTCCCAAGAGTGGCCCGAGTGTGCGGGGAGCTTGGGGG 927
Db 817 TTTGTGTTCAGGGTGTGATCCCAAGATAGCAAGAGTGTCAAGGGGAGCCAGGGCGG 876
Qy 928 AGGAAGACCTTCAGCAGAGATGGACAGCTTTTCTGAAGCTGACCTGCTGCTCCAGGG 987
Db 877 CTGAGGACCTTGCAAGAATAATGGACCTCTTCTTGAAGCCCGGACTCATCTGCTCCCG 936
Qy 988 CCGAGCATGGCCGGGCTCCGGGTTCTCAGGCTATGGCAGACTTTCGGCTCAGCTC 1047
Db 937 CCAGACAGCGCTTGGTCTTCAATGTCTGGGATGCTTCTGCTCAGCTC---CCCG 993
Qy 1048 GGAGCGGAACCCCATCTTTTATGGATCTTTTCTCCAGATGGGAGGAGCTGCGCATC 1107
Db 994 GGCTGAAGGTGCTGTGTCTATGCACTTTCACCCCAAGCTGTAACAAGTGGGGTG 1053
Qy 1108 TCTGCTGTGTGCC-----TTCCGACCCCAAGACATCCGGGAGTGTCTCAATGGT 1158
Db 1054 TCGCAGTGTGCGCTTACAACCTGTCCACAGCGGAGGAGTCTTCTCCCGGGAAGTAC 1113
Qy 1159 CCCTTTAGAGAGCTAAACATGACTGCAACAGGGGAGTGTCTGTCATGGACAAGAGTG 1218
Db 1114 ATGACAGACCAAGTGGAGCAGTCCACACCAAGTGGGTGGCTATATGGCCCGGTA 1173
Qy 1219 CCCCAGCCAGACCTGGAGAGTGCATCGCCCAACAAATGAAGCTCCAGCAGTGTGGATCC 1278

Db	1174	CCCAAGCCGGGCTTGAGGGTGCATGCAGCGAGGCGACGGCCGCCAACTACACCAGC	12338
Qy	1279	TCACTCTCCCTGCAGCACCGCTGCACCTTTATCAGACACCACTCTCATGTGCACAGG	1338
Db	1234	TCCTTGAAATTTGCCAGACAAGCGCTGCATGTCGTTAAAGACCACTTTGATGGATGAC	1293
Qy	1339	CCCGTGTTCCGGCTGACGCCGCCGCCCTCTCGTCTACTACAGATACAGCCTATCTCAGA	1398
Db	1294	TCGGTAACCCCAATAGACAACAGGCCCGAGGTTAATCAAGAAGATGTGAACCTACACCCAG	1353
Qy	1399	GTGGTGGCCACAGGGTGACCAAGCCTCTCAGGGAAGAATATGACGTGCTCTACCTGGGG	1458
Db	1354	ATCGTGGTGACCGGACCCAGGCCCTGGATGGGACTGCTATGATGTCATGTTGTGCAGC	1413
Qy	1459	ACAGGATGGACACACTCCACCGGGTGTGGCGATTGGAGCTCAGCTCAGTGTCTTGGAG	1518
Db	1414	ACAGACGGGAGCTCTGCACAAGCCATCAGCCTCGAGGACGCTTTCACATCATCGAG	1473
Qy	1519	GATCTGCGCTTGTTCAGAACACACACCGCGTTGAGAGCATGAATTTGTA-----	1568
Db	1474	GAGACCCAGCTCTCCAGGACTTTGAGCCAGTCCAGACCTGCTGCTTCAAGAAG	1533
Qy	1569	--CCAGATTGGCTCTCGTGGGCTCCCATACTGAGGTGACACAAGTGAACACCAAC	1626
Db	1534	GGCAACAGGTTGTGCTATGCTGGCTTAACTCGGGGCTGTCCAGGCCCGCTGGCGCTTC	1593
Qy	1627	TGTGGCCGCTCTCCAGAGCTGCTCGGAGTGATCTTGCCCCAGGACCCCGTGTGCGCTTG	1686
Db	1594	TGTGGGAAGCACGCACTCTCGAGGACTGTGCTGCGCGGGACCCCTACTGCGCTG	1653
Qy	1687	AGCTTCGGGCTTGATGCTTGTGGGCC	1714
Db	1654	AGCCGCCACACGCACTCGTGGGTC	1681

RESULT 10

```

PCT-US02-36071A-165
; Sequence 165, Application PC/TUSO236071A
; GENERAL INFORMATION:
; APPLICANT: Sagres Discovery
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000142
; CURRENT APPLICATION NUMBER: PCT/US02/36071A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: US 10/052,482
; PRIOR FILING DATE: 2001-11-08
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 165
; LENGTH: 2586
; TYPE: DNA
; ORGANISM: Mus musculus
PCT-US02-36071A-165

```

[illegible]

Qy	429	GAAGAAGGCAA----	GAAGAGGAGCAAGATGTCA	CAAAATTTTAT	TCAGATTCT	CGCCATGT	485
Db	294	AGAGAGGGGAAATCA	AAAGCAGACGGAATGCCT	AAACTACAT	TCAGTTCG	AGTACAGCCACT	353
Qy	486	CAATGCCTCTCACCT	CTCAGTCGCGCACCTT	CGCTTTGAT	CCGAAGTC	CGGGTTAT	545
Db	354	AAGCAGACATTTCC	CTATGTGTGGACCAAT	GCCTTCAG	CCCACTCT	GTGACCACCT	413
Qy	546	TGATGTGTCCAGTT	TCACAGAGTTGA	AAAGACTTGA	GAGCGCGCGGGG	AAATTCCTCTTT	605
Db	414	GAACCTTGACATCT	TCAGTTTCGGGAA	AAAGTGAAGATG	GCAAAAGGA	AGATCCCCCTT	473
Qy	606	TGAGCCAGCTCAAC	GGTCAGAGCTGAAT	GGCTGGGGGGCT	CTCTACAC	CGCCACTGT	665
Db	474	CGACCCGCCACAG	CTACACATCAGT	CTACGTTGGGG	CGAGCTTACT	CTCGGAGCTC	533
Qy	566	GAAGAACTTCTGGG	AGCTGAGCCCATCAT	CTCCGAGCTGT	GGGTCCGAGCT	GAGGACTG	725
Db	534	CTATAATTTCTTG	GGCAGTGAACCCAT	CATCTCTCG	AAACTCTTCC	CACAGTCCCTT	594
Qy	726	GATTCGAACAGAG	ACCTTGTTCATCT	CGCTTAAATGCT	CCAGCCCTTGT	CGCAGCTATGGT	785
Db	594	G-----ACGAG	TATGCCATCCGCT	GGCTGAAC	AGCGCTTCTT	CTTGTCTGCTGAGCT	647
Qy	786	CCTGAGCCCCAG	CTGAGTGGGGGAT	GAAGATGG	AGACGATGA	AAATCTTTTTTTTCTT	845
Db	648	GATCCAGAAAAG	CCAGATGGTCCG	GAGGGTGAAGAT	GACAAGCT	CTACTTCTTTTTTAT	707
Qy	846	GGAGACCTCCG	AGTGTGACTCTAT	GAGCGCATCA	AGGTCCCA	AGAGTGGCCCGAGT	905
Db	708	GGAGGTATCCG	TGGAGTACGAA	TTTCTTCA	AGTTGATGAT	CCCGAGTGTGCCAGGGT	767
Qy	906	GTGTGCGGGG	ACCTTGGGGCG	AGAAAGCCCTT	CAGCAGAGAT	GGAGAGCTTCTTGAA	965
Db	768	GTGCAAGGGC	CAACAGGGCGCT	CGGACTTTG	CAAAAAGT	TGAGCTCTTCTTAA	827
Qy	966	GGCTGAGCTGT	GTGCCAGGGCC	AGCAGTGG	CGGGGCTC	CGAGGCTAT	1025
Db	828	GGCAGGCTGAT	CTGCTCCAAG	CCACAGACAGT	GGCGCTGGT	CTTCAACATACTT	887
Qy	1026	GGCAGAGTTCG	GCGCTCAGCCT	CGAGCGGNA	CCCCCATCTTT	TATGGGATCTTTT	1085
Db	888	GTTTGTCTGAG	GGG-----	CCCGGGCTCA	AGGAGCGCTGT	GTCTATCGCGTCTT	944
Qy	1086	CCAGTGGAGG	AGCTGCATCT	CTGTCTGTGT	GGCTTCG	AGCCCCCAAGACAT	1145
Db	945	ACACTGAACNA	TGTGGTCTGT	CAGGGGTGT	CGCCTAC	ACTGCCACGGTGGAGC	1004
Qy	1146	AGT-----GCT	GAATGGTCCCTTT	AGAGAGCTAAA--	ACATGACT	GTCAACAGGGGACT	1196
Db	1005	AGTCTTCTCCG	TGGAAAGTACAT	GCAGAGTGC	ACAGTGG	AGCAGTCTCA	1064
Qy	1197	GCCTGTCTATG	GACAAACAGG	TGCCACAG	CCAGACCTT	GGAGAGTGCAT	1256
Db	1065	GGTCCGCTACA	ATGGCCCA	GTCAGCTCC	CCCGACCTT	GGAGCGTGTAT	1124
Qy	1257	GAAGCTCCAG	CAGTTTGGATCT	CACTCTCC	CTCCAG	CGCGTGTCTCACTTTTAT	1316
Db	1125	CCGGGAGCCNA	CTTACACCA	GCTCTTGAATCT	CCCCAG	CAAAACACTGCAGTTTGT	1184
Qy	1317	AGACCACTCT	CTATGACAG	CGCCGCTGTT	CCCGGCTG	AGCGCCGCCCTCTG	1376
Db	1185	AGACCACCTT	TGATGATGACT	CTAGTGA	CCCCGATAG	ACAACAGACCC	1244
Qy	1377	TACAGATAC	AGCCCTATCT	CAGATCTG	TGGCCCA	AGGAGTGTAC	1436
Db	1245	AAAAGATGT	AAACTTAC	CCCAAGATAG	TGTGAG	ACAGGACCC	1304
Qy	1437	ATATGAGCT	GCTCTACCT	TGGGACAG	AGGATGG	ACACTCC	1496
Db	1305	CTACGAGCT	CAATGTTCA	TACAGCA	CACCGGG	AGCTCTG	1364
Qy	1497	AGCTCAGCT	CAGTGTCT	TGGAGGATCT	TGGCTTGT	CCCCAGAAC	1556

Db 1365 AAAGAGGTGCATGTCATCAGGAGAGACCAACTCTTCGGGACTTTGAACCGGTCTCTAAC 1424
Qy 1557 CATG-----AAATTGTACCACGATTGGCTCTCTGGTGGCTCCCATCTAGAGT 1604
Db 1425 TCTGCTCTATCGTCAAGAAGGGGAGGAAGTTTGTCTATGCAAGCTTCAACTCTGGAGT 1484
Qy 1605 GACACAAAGTGAACACCAACAGCAACTGTGGCCGTCTCCAGAGCTGTCTGGAGTGTATCTCTGGC 1664
Db 1485 GGTCAAGCGCCCTGGCAATTCGCGAAAGACAGCGGTAGCTGTGAAGACTGTGTGTAGC 1544
Qy 1665 CAGGAGCCCGTGGCGCTGGAGCTTCGGCTTGTGATGCTGTGTGGCC 1714
Db 1545 ACGGAGCCCTACTGTGCTGGAGCCAGCCATCAAGGCGCTGTGTACCC 1594

RESULT 11
PCT-US02-36071A-164
; Sequence 164, Application PC/TUS0236071A
; GENERAL INFORMATION:
; APPLICANT: Sagres Discovery
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000142
; CURRENT APPLICATION NUMBER: PCT/US02/36071A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: US 10/052,482
; PRIOR FILING DATE: 2001-11-08
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 164
; LENGTH: 2769
; TYPE: DNA
; ORGANISM: Mus musculus
PCT-US02-36071A-164

Query Match 5.1%; Score 205.6; DB 1; Length 2769;
Best Local Similarity 49.2%; Pred. No. 1e-44;
Matches 733; Conservative 0; Mismatches 724; Indels 33; Gaps 6;

Qy 249 TCTCACCCGGTTTCACGGTCTCATAGTACAAATTAATCTGCTCTCTCTTGGATCTGCG 308
Db 174 TCTGGTGCAGTTTTCACAAAGCCAGGCATCTTTAACTACTCGGCGCTTGTGATGAGGA 233
Qy 309 CTCCACACACTTTTACCTGCGTGCACGGATAGCATCTTGGCTTTAACCTTCCCTTCTC 368
Db 234 CAAGACACTCTGTATGTAGCGCCCGGGAAGCAAGTCTTTGCACTGAATCGCTTGAACAT 293
Qy 369 TGGGGAAGACCCCGAAGGATCGACTGGATGGTACCTGAGACTCACAGACAGAAGTCCAG 428
Db 294 CTCTGAGAAGCAACATGAGGTATATTGGAAGTCTCTGAAGACAAAAATCCAAAGTGTG 353
Qy 429 GAAGAAAGGCAA----GAAGAGGACGAATGTCAAAATTTATCCAGATTTCTCGCCATGT 485
Db 354 AGAAGAGGGGAAATCAAGACAGAGCGGAATGCTTAACTACATTCGAGTACTACAGCCACT 413
Qy 486 CAATGCTCTCACCTCTCAGTGCAGCTCGCCACTTTCGCTTTTGCATCCGAAGTCCGGGTAT 545
Db 414 AAGCAGCACTTCCCTCTATGTGTGGGACCAATGCGTTCCAGCCCACTGTGACCACCT 473
Qy 546 TGATGTCTCCAGTTTCCAGCAGGTTGAAAGACTTTGAGAGCGCGCGGGGAAATCTCCTTT 605
Db 474 GAATCTGACATCTTCAAGTTTCTGGGAAAGATGAAGATGGCAAGAGATGCCCTT 533
Qy 606 TGAGCCAGCTCAACGGTGCAGAGCTGTAATGGCTGGGGGGTCTCTACACCCGCCACTGT 665
Db 534 CGACCCCGCCACAGCTACACATCAGTCAATGTTGGGGGCGAGCTCTACTCTGGCAGCTC 593
Qy 666 GAAGAACTTCTGGGAGTACAGCCCATCATCTCCGAGCTGTGGGTGCAGCTGAGGACTG 725
Db 594 CTATAATTTCTTGGGAGTGAACCCCATCATCTCTCGAAACTCTTCCACACAGTCCCTTGAG 653

Qy 726 GATTGGAACAGAGACACTTGTCTATCCTGGCTTAATAGTCTCCAGCCCTTTCTCGACGCTATGGT 785
Db 654 G-----ACGGAGTATGCCATCCGCTGGCTGAACGAGCCTAGCTTCTGCTTGTGCTGACGT 707
Qy 786 CTTGAGCCCACTGAGTGGGGGATGAAGATGGAGACGATGAATATCTTTTTTCTTCAC 845
Db 708 GATCCAGAAAAGCCAGATGGTCGGAGGGTGAAGATGAAGAAGTCTACTCTTTTTTAC 767
Qy 846 GGAGACCTCCCGAGTGTGACTCTATGAGCGCATCAAGGTCCCAAGAGTGGCCCGGT 905
Db 768 GGAGTATCCGTGGATACGAATTCGTTCAAGTTGATGATCCCGGAGTGGCCAGGT 827
Qy 906 GTGTGCGGGGACCTTGGGGGACAGAAACCTTCAGCAGAGATGGAGAGCTTTCTGAA 965
Db 828 GTGAAGGGGACACAGGGCGCTTCGGACTTTGCAAAAAAAGTGGACCTCTTCTCTTAA 887
Qy 966 GGCTGACCTGCTGTGCCCCAGGCGCCGAGCATGGCCGGCCCTCCGGGGTCTTCGAGGCTAT 1025
Db 888 GGCCAGGCTGATCTGCTTCCAAAGCCAGACAGTGGCTGTCTTCAACATACTTCAAGATGT 947
Qy 1026 GGCAAGCTTCGGCTCAGCTTGAGCGGGAACCCCATCTTTTATGGGATCTTTTCTC 1085
Db 948 GTTTGTGCTGAGGC---CCGGGCTCAAGGAGCTGTGTTCTATGGGTCTTCAACCC 1004
Qy 1086 CCAGTGGGAAGAGCTGCCATCTCTGTGTGTGTGCTTCGACCCCAAGACATCCGGGC 1145
Db 1005 ACAGCTGAACAAATGTGGTCTCTCAGCGGTGTGGCTTACACACTGGCCACGGTGGAGC 1064
Qy 1146 AGT-----GCTCAATGGTCCCTTTAGAGAGCTTAA--ACATGACTGCACACAGGGNACT 1196
Db 1065 AGTCTTCTCCCGTGGAAAGTACATGCAAGTGGCCACAGTGGAGCAGTCTCACACCAAGTG 1124
Qy 1197 GCCTGTCTATGACAACAGAGTGGCCCGCCAGACCTTGGAGAGTGCATCGCCCAACACAT 1256
Db 1125 GGTGCGTACAAATGGCCAGTGGCCCACTCCCGGAGTGGAGCTGTATCGACAGTGGGC 1184
Qy 1257 GAAGCTCCAGCAGTTTGGATCTCTCACTCTCCCTGCCAGACGCGTGTCTACCTTTATCAG 1316
Db 1185 CCGGCGACCAACTACACCACTCTTGAATCTCCAGACAAAACACTGCAAGTTGTAAA 1244
Qy 1317 AGACCACTCTCATGACAGGCGCGTGTCCCGCTGACGGCGCCCTGCTGGTGCAC 1376
Db 1245 AGACCACTCTTGTATGATGATGACTCAGTGACCCCGATAGACAAACAGCCCAAGCTGATCA 1304
Qy 1377 TACAGATACAGCTCTCTCAGAGTGTGGCCCAAGGCTGACAGCTCTTCAGGGAAAGA 1436
Db 1305 AAAAGATGTAACACTACACCAAGATAGTGTAGACAGGACCCAGCCCTGGATGGACTTT 1364
Qy 1437 ATATGAGCTGCTTACCTGGGACAGAGATGGACACTTCCACCGGGGTGTGCGCATGG 1496
Db 1365 CTAGCAGCTCATGTTTATCAGCAGACAGCGGGAGCTCTGCTATAAGCAGTCACTCTTAC 1424
Qy 1497 AGCTCAGCTCAGTGTCTTGGAGATCTGGCTTTGTTCCAGAACACAGCCGGTGTAGAG 1556
Db 1425 AAAAGAGGTGATGCTCATCTCAGAGAGACCAACTCTTCCGGGACTTTGAACCGGCTTAAC 1484
Qy 1557 CATG-----AAATTGTACCAGTGTGCTCTCTGTGGCTCCCTACTACTGAGT 1604
Db 1485 TCTCTGCTGATGCTCAAGAAGGGGAGGAAGTTGTCTATGAGGCTCCCACTCTGAGT 1544
Qy 1605 GACACAAAGTGAACACCAAGCACTGTGGCCGTCTCCAGAGCTGCTCGGAGTGTATCTCTGGC 1664
Db 1545 GGTCAAGCGCCCTGGCATTTCTCGAAAAGCAGCGTAGCTGTGAAGACTGTGTGTAGC 1604
Qy 1665 CCAGAGCCCTGTGGCGCTGGAGCTTCCGGCTTGTGCTGTGTGGCC 1714
Db 1605 ACGGAGCCCTTACTGTGCTGGAGCCAGCCATCAAGGCGCTGTGTGTACCC 1654

RESULT 12

PCT-US03-27411-44

; Sequence 44, Application PC/TUS0327411

; GENERAL INFORMATION:

THIS PAGE BLANK (USPTO)

Result No.	Query #			ID	Description	
	Score	Match	Length			
1	1673.6	41.8	2893	11	US-09-808-665A-1	Sequence 1, Appli
2	1673.6	41.8	2894	11	US-09-808-665A-3	Sequence 3, Appli
3	654	16.3	799	11	US-09-808-665A-13	Sequence 3, Appli
4	597.8	14.9	928	14	US-10-188-246-13	Sequence 13, Appli
5	527.4	13.2	777	14	US-10-188-246-15	Sequence 15, Appli
6	251.8	6.3	3293	14	US-10-149-819-25	Sequence 25, Appli
7	248.6	6.2	2155	13	US-10-003-152-11	Sequence 11, Appli
8	248.6	6.2	2155	14	US-10-002-050-11	Sequence 11, Appli
9	248.6	6.2	2155	14	US-10-002-304-11	Sequence 11, Appli
10	248.6	6.2	2156	13	US-10-003-152-21	Sequence 21, Appli
11	248.6	6.2	2156	14	US-10-002-050-21	Sequence 21, Appli
12	248.6	6.2	2156	14	US-10-002-304-21	Sequence 21, Appli
13	248.6	6.2	2284	13	US-10-003-152-13	Sequence 13, Appli
14	248.6	6.2	2284	14	US-10-002-050-13	Sequence 13, Appli
15	248.6	6.2	2284	14	US-10-002-304-13	Sequence 13, Appli
16	247	6.2	3781	11	US-09-946-374-252	Sequence 252, App

Db 2198 CTGGCAAGAGGGCAGTGGGTTTTGGTGGATTCTCACCACCCCTTCTGCTTGATCCTTGC 2257
Qy 2341 CCAAGCCAGCCACATCCGCTCACTGGGGGCTCTAGCCAGCTGTGATGAGACCTCC 2400
Db 2258 CCAAGCCAGCCACATCCGCTCACTGGGGTCTCTAGCCACATGTGATGAACATCC 2317
Qy 2401 ATCTAAAGCCGGGAAATAGCTGCCAGCCATGAGCAGCTCTGTGAAAC 2448
Db 2318 ATCTAGAGCTGGCAATGACACACTAGTGTATAGTGATCAGTGAAC 2365

RESULT 2

US-09-808-665A-3
; Sequence 3, Application US/09808665A
; Publication No. US2003002282A1
; GENERAL INFORMATION:
; APPLICANT: David Michalovich
; APPLICANT: Trudy Rachel Doe
; APPLICANT: Philip David Hayes
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP-30036-C1
; CURRENT APPLICATION NUMBER: US/09/808,665A
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 09/160,762
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 98300693.3
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 9816423.9
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: 9816676.2
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2894
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-808-665A-3

Query Match 41.8%; Score 1673.6; DB 11; Length 2894;
Best Local Similarity 82.2%; Pred. No. 0;
Matches 2012; Conservative 0; Mismatches 319; Indels 117; Gaps 3;
Qy 1 GCCAGGCGCCGCGAGTAGTACTAGTAGAGGCTGCTGAGCGCGCCACCCGCGCAC 60
Db 35 GCCAGGCGCCAGTAGCCCGGGCCCTGAGAGAGCCCTAGCTTGGCGCGCACCCGCGC 94
Qy 61 CAGGCGGAGCAGAGATGCTTGGCCAGGCGCGAGGCGCCCGCGGGCGCCCGCGCGCT 120
Db 95 CAGGCGGAGCAAGATGCGGCGCTCTGCTGGCGGCGCCCGCGGGTCCCGGCGAGCT 154
Qy 121 CCGGTCTTCCCTTCCCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Db 155 ACAGCCTCGCCCTTCCC-----GCTACTGCTGCTGGCGGCTGCTGAGCGGC 199
Qy 181 CCGGTGTCGGCGCGCGTCCCGCGCTCACTGAGCGCAGAACCTCGCTGCCATCTCCGAGGCT 240
Db 200 CCGGTATCCGGCGCGTCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 259
Qy 241 GACTCTCTATCACCAGGTTTGCAGCGCTCTCATAGTACAATTAATCTGCTCTCTCTGTG 300
Db 260 GACTCTCTATCACCAGGTTTGCAGCGCTCTCATAGTACAATTAATCTGCTCTCTCTGTG 319
Qy 301 GATCCTGCTCCACACACTTACGTCGGTGCAGGGATAGCATCTTCGCTTTTAAACCTC 360
Db 320 GATCCTGCTCCACACACTTATGTTGGCGCGCGGACACCATCTTCGCTTTTATCCCTG 379
Qy 361 CCCTTCTCTGGGAAAGCCCGAGGATGACTGGATGGTACCTTGAGACTCAGACAG 420
Db 380 CCCTTCTCAGGGAGAGACCCCGAGGATGACTGGATGGTTCCTGAGGCTCAGACAG 439
Qy 421 AACTGCAGGAAGAGCAAGAAAGAGGACGAATGTCAATAATTTATCCAGATTCTCGCC 480

Db 440 AACTGTAGGAAGAAAGCAAGAG----- 466
Qy 481 ATTGTCAATGCTCTCACCTCCTCAGCTGCGGCACCTTCGCTTTTGATCCGAAGTGGCGG 540
Db 467 -----GATGTGTCAGGTTCCAGAGGTTCAAGACTTGAGAGTGGCGGGGAAATGT 466
Qy 541 GTTATTGATGTCCAGTTTCCAGCAGGTTGAAAGACTTGAGAGCGCGCGGGGAAATGT 600
Db 467 -----GATGTGTCAGGTTCCAGAGGTTCAAGACTTGAGAGTGGCGGGGAAATGT 520
Qy 601 CCTTTTGAGCAGCTCAACGGTTCAGCAGCTGTAAATGCTGGGGGCTCTCTTACACCGCC 660
Db 521 CCTTTTGAGCAGCTCAGCGTTCAGCAGCTGTAAATGCTGGGGGCTCTCTATGCTGCC 580
Qy 661 ACTGTGAAGAACTCTCTGGGACTGAGCCCATCATCTCCGAGCTGTGGGTGCGAGCTGAG 720
Db 581 ACTGTGAAGAACTACCTGGGACGAGCCAAATATCACAGAGAGTGGGTGCTGCCGAG 640
Qy 721 GACTGGATTGGAACAGAGACCTTGTATCTCTGGCTTAATGCTCCAGCCCTTTGTTCAGCT 780
Db 641 GACTGGATTGGAACAGATACCTTGCCTTCTCTGGCTGAACGCCCCAGCCCTTTGTTCGAGCC 700
Qy 781 ATGTCTCTGAGCCAGCTGAGTGGGGGATGAAGATGAGAGATGAATCTTTTTC 840
Db 701 GTGGCCTTGAGCCAGCCGAATGGGGGATGAAGATGAGAGACGAGAAATCTACTTCTTC 760
Qy 841 TTCAGGAGAGCTCCCGAGTGTGGACTCTTATGAGCCCATCAAGGTCCCAAGAGTGGCC 900
Db 761 TTTACGGAGACTTCCCGAGCATTTGACTATACAGAGCCATTAAGTCTCCAGCGGTGCC 820
Qy 901 CGAGTGTGTGGGGGAGCTTTGGGGGAGAGAGACCTTTCAGCAGAGATGAGCAGCTTT 960
Db 821 CGTGTGTGTGGGGGAGCTTGGGGGCGGAAGACCTTCAGCAGAGATGAGCAGCTTT 880
Qy 961 CTGAAGCTGACCTGCTGTGCCAGGCGCCGAGCATGCGCGGCTCCCGGGTCTCTGAG 1020
Db 881 TTGAAGCTGACCTGCTGTGCCAGGCGCTGAGCATGCGCGGCTCCAGCTGTCTCTGAG 940
Qy 1021 GCTATGCGAGAGCTTCGCGCTCAGCTTGAGCGGGAACCCCATCTTTATGCGGATCTTT 1080
Db 941 GATGTGTGCTGCTGCTGAGCTTGGGCGAGGAGTCCCATCTTTTATGCGCATCTTT 1000
Qy 1081 TCTCTCCAGTGGGAGAGCTGCCATCTGCTGTGTGCTGCTTCCGACCCCAAGACATC 1140
Db 1001 TCTTCCAGTGGGAGGCTACTATCTGCTGCTGCTGCTTCCGACCAAGACATTT 1060
Qy 1141 CGGCGAGTGTGAATGTCTCTTTAGAGAGCTAAACATGACTGCAACAGGGAGCTGCT 1200
Db 1061 CGGCGAGTGTGAATGTCTCTTTCAGAGAACTTAAACATGACTGCAACAGAGGAGCT 1120
Qy 1201 GTCATGGACAGAGGTCGCCAGCCAGCTGGAGAGTGCATCGCCCAACAACTAAG 1260
Db 1121 GTCTGGACAATGATGTGCCCGAGCCAGAGCTGGAGAGTGCATCACCACAACTAAG 1180
Qy 1261 CTCAGCAGATTTGGATCTCTACTCTCCCTGCCAGACCCGCTGCTCACCTTTATCAGAG 1320
Db 1181 CTCGGCATTGCTGCTCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1240
Qy 1321 CACCTCTCATGAGAGCGCTGTTCCCGCTGACGGCGCGCCCTGCTGCTGCTGCTGCTG 1380
Db 1241 CACCCATCATGAGAGCGCAGTGTTCAGCTGATGGCCACCCCTGCTGCTGCTGCTGCTG 1300
Qy 1381 GATACAGCTATCTCAGAGTGTGGCCCAACAGGCTGACAGCCCTCAGGGAAGAATAT 1440
Db 1301 GATACAGCTATCTCAGAGTGTGGCCCAACAGGCTGACAGCCCTCAGGGAAGAATAT 1360
Qy 1441 GAGCTGCTTACCTGGGAGAGAGATGACACCTTCCACCCGGCTGTGCGGATTTGAGCT 1500
Db 1361 GATGTCTCTACCTGGGAGAGAGATGACACCTTCCACCCAGAGCTGCGGATGCGAGCT 1420
Qy 1501 CAGCTCAGTGTCTTGGAGGATCTGGCTTTGTTCAGAGAACACAGCCGCTTGGAGAGATG 1560
Db 1420 CAGCTCAGTGTCTTGGAGGATCTGGCTTTGTTCAGAGAACACAGCCGCTTGGAGAGATG 1560

Db 1421 CAGCTCAGCGTTCTTGAAGATCTGGCCCTTATTCACAGAGCCAGCCAGCTTGAGAACATG 1480
QY 1561 AAATTGTACACAGATTGGCTCTCTGGTGGGCTCCCATCTAGAGGTGACACAAGTGAACACC 1620
Db 1481 AAATTGTACACAGCTGGCTCTCTGGTGGGCTCCCATCTAGAGGTGACACAAGTGAATACA 1540
QY 1621 AGCAACTGTGGCGCTCTCAGAGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCT 1680
Db 1541 ACCAAGTGTGGCGCTCTCAGAGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCT 1600
QY 1681 GCTTGAGCTCTCGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCT 1740
Db 1601 GCTTGAGCTCTCGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCT 1660
QY 1741 CAAGATATAGAGTACAGGATCT 1800
Db 1661 CAAGACATAGAGTACAGGATCT 1720
QY 1801 GTAGTGTGAAGTTCGGTGGCT 1860
Db 1721 GTAGTGTGAAGTTCGGTGGCT 1780
QY 1861 TCTGCTGGGCT 1920
Db 1781 TCAGCATGGGCT 1840
QY 1921 AGGATGGACTAGAGTGGTGTGACCCAGGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1980
Db 1841 CGGGATGGACTAGAGTGGTGTGACCCAGGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1900
QY 1981 CAGGAGGTGGAGCCCGCT 2040
Db 1901 CAGGAGGTGGAGCCCGCT 1960
QY 2041 GGACCTCTAAACCGGGCCACACCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100
Db 1961 GATGCTCCAGCGGGGCCACAC--AGTGGGGGGGGGCTCTCTCTCTCTCTCTCTCTCTCT 2017
QY 2101 GTTCTTGGAGCATCCCTCACT 2160
Db 2018 ATTCTCGAGCATCCCTGACT 2077
QY 2161 AGGGAGCTCTAGTAGACACAGGTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2220
Db 2078 AGGGAATCTTGCTAGACAAAGTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2137
QY 2221 ACAAGCTATAGTCAGGACCT 2280
Db 2138 ACNAGCTAGACCAAGACCT 2197
QY 2281 CTGGGTAAGCGGGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2340
Db 2198 CTGGGCAAGAGGGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2257
QY 2341 CCNAGCCAGCCACATCCGGCTCACTGGGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2400
Db 2258 CCNAGCCAGCCACATCCGGCTCACTGGGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2317
QY 2401 ATCTAAGCGGGGAAATGACTGCCAGCATGAGCAGTCTCTGGAAC 2448
Db 2318 ATCTAGAGTGGGCAATGACCACTAGTGTATAGTGATCACTGGAAAC 2365

RESULT 3
US-09-808-665A-5
; Sequence 5, Application us/09080665A
; Publication No. US2003002282A1
; GENERAL INFORMATION:
; APPLICANT: David Michaelovich
; APPLICANT: Trudy Rachel Doe
; APPLICANT: Philip David Hayes
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP-30036-C1

; CURRENT APPLICATION NUMBER: US/09/808, 665A
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 09/160,762
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 98300693.3
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 9816423.9
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: 9816676.2
; PRIOR FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 799
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (720)
; OTHER INFORMATION: OTHER INFORMATION: Partial EST derived sequence
US-09-808-665A-5

Query Match 16.3%; Score 654; DB 11; Length 799;
Best Local Similarity 88.6%; Pred. No. 2e-170;
Matches 708; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 1126 CGACCCCAAGACATCCGGCAGTGTCTGAATGGTCCCTTTAGAGAGCTAAACATGACTGC 1185
Db 1 CGACCACAGACATTCGGACAGTGTGATGGTCCCTTCAGAGAACTAAACATGACTGC 60

QY 1186 AACAGGGGACTGCCTGTCTATGGACACAGAGGTGCCGCCAGCCAGACCTGGAGAGTGCATC 1245
Db 61 AACAGAGGACTGCCTGTCTGTGGACATGATGTGCCCCAGCCAGACCTGGAGAGTGCATC 120

QY 1246 GCCAACAACTAGAGTCCAGCAGTTGGATCTCTCACTCTCCCTGCCAGCCGCGTCTC 1305
Db 121 ACCAACAACTAGAGTCCGGCAGTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180

QY 1306 ACCTTTATCAGAGACACACCTCTCTATGGAGAGCCGCTGTCCCGGCTGACGGCCGCCCC 1365
Db 181 ACCTTCATCCGGACACACCTCTCTATGGAGAGCCGCTGTTCACAGCTCATGGCCACCCC 240

QY 1366 CTGCTGGTCACTACAGATACAGCTATCTCAGAGTGTGGGCCACAGGGGTGACAGCCCTC 1425
Db 241 CTGCTGGTCACTACAGTTCAGAGCTATCTCAGAGTGTGGGCCACAGGGGTGACAGCCCTC 300

QY 1426 TCAGGGAAGATATAGCTGTCTACCTGGGACAGAGATGGACACCTCCACCGGGCT 1485
Db 301 TCAGGGAAGATATAGTGTCTACCTGGGACAGAGATGGACACCTCCACCGAGCA 360

QY 1486 GTGCGCATTTGGAGCTCAGCTCAGTGTCTTTGGAGGATCTGGCCTTGTTCACGAACACACAG 1545
Db 361 GTGCGGATCGGAGCTCAGCTCAGCTCAGCTTCTTGAAGATCTGGCCTTATTCACGAGCCACAG 420

QY 1546 CCGGTTGAGAGCATGAAATTTGACACGATTTGGCTCTCTGGTGGGTCCCATCTAGAGTG 1605
Db 421 CCAGTTGAGAACATGAAATTTGACACAGTGGCTCTCTGGTGGGTCCCGTACTGAGGTG 480

QY 1606 ACACAAGTCAACACAGCAACTGTGGCGCTCTCCAGAGCTGTCCGAGTGTATCTCTGGCC 1665
Db 481 ACACAAGTCAACACCAACTGTGGCGCTCTCCAGAGCTGTCCGAGTGTATCTCTGGCC 540

QY 1666 CAGGACCCCGTGTGGCCTTGGAGCTTCCGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1725
Db 541 CAGGACCCAGTCTGTGGCTGGAGCTTCCGGCTGGATGAGTGTGTGTGTGTGTGTGTGTGTGTGT 600

QY 1726 CACCGGGGATGGTTCAAGATATAGAGTACAGCGGATGTCTCTCTTTGTGTGTGTGTGTGTGTGT 1785
Db 601 CACCGGGGTTGGTCCAAAGACATAGAGTACAGAGATGTCTCTCTTTGTGTGTGTGTGTGTGTGT 660

QY 1786 CCTGGAGAACATCCCGTAGTGTGAAGTTCCGGTGGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1845
Db 661 CCTGGAGAACGTCCAGTAGTGTGAAGTTCCCGTGGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGT 720

;; PRIOR APPLICATION NUMBER: 60/304502
;; PRIOR FILING DATE: 2001-07-11
;; PRIOR APPLICATION NUMBER: 60/305011
;; PRIOR FILING DATE: 2001-07-12
;; PRIOR APPLICATION NUMBER: 60/306085
;; PRIOR FILING DATE: 2001-07-17
;; PRIOR APPLICATION NUMBER: 60/326981
;; PRIOR FILING DATE: 2001-10-14
;; PRIOR APPLICATION NUMBER: 60/360923
;; PRIOR FILING DATE: 2002-02-28
;; PRIOR APPLICATION NUMBER: 60/363636
;; PRIOR FILING DATE: 2002-03-12
;; PRIOR APPLICATION NUMBER: 60/373063
;; PRIOR FILING DATE: 2002-04-16
;; NUMBER OF SEQ ID NOS: 60
;; SOFTWARE: Custom
;; SEQ ID NO 15
;; LENGTH: 777
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (1)..(777)
US-10-188-246-15

Query Match 13.2%; Score 527.4; DB 14; Length 777;
Best Local Similarity 90.2%; Pred. No. 2.2e-135;
Matches 564; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 193 CCGCTCCCGCTAGTGGCCGAGAACCTCGCTGCCATCTCGAGGCTGACTCTCTATCTC 252
DB 7 CCGCTCCCGCTAGTGGCCGAGAACCTCGCTGCCATCTCTGAGGCTGACTCTCTATCTC 66
QY 253 ACCCGGTTTGACGCTGTCATACGTTACATTTACTGCTCTCTCTGTCGATCTCGCTCC 312
DB 67 ACCCGGTTTGACGCTGTCATACGTTACATTTACTGCTCTCTCTGTCGATCTCGCTCC 126
QY 313 CACACACTTTAGCTCGGTCACGGGATAGCATTTTCGGCTTTAACCTCCCTCTCTCTGGG 372
DB 127 CACACACTTTATGTTGGCGCGGACACCATCTTCGCTTTATCTCTCTCTCTCTCGG 186
QY 373 GAAAGACCCGAGGATGAGTGGATGTTACTGAGCTACAGCTACAGACAGAACTGCAAGG 432
DB 187 GAGAGACCCGAGGATGAGTGGATGTTCTGAGGCTACAGACAGAACTGTAGGAG 246
QY 433 AAAGCAAGAAAGAGAGGAATGTACAAATTTTATCCAGATCTCGCCATTTCTCAATGCC 492
DB 247 AAAGCAAGAAAGAGAGGAATGTACAAATTTTGTCCAGATTTCTGCCCATTTGCCAATGCC 306
QY 493 TCTACCTCTCTACGTCGGGACCTTCGCTTTTGTATCCGAGTGTGGGGGTTATGATGTG 552
DB 307 TCTACCTCTCTACGTCGGGACCTTCGCTTTTGTATCCGAGTGTGGGGGTTATGATGTG 366
QY 553 TCCAGTTTCCAGCAGTTGAAGACTTGAGCGCGCGGGGGAATGTCTTTTTCAGCCA 612
DB 367 TCCAGTTTCCAGCAGTTGAAGACTTGAGAGTGGCGGGGGAATGTCTTTTTCAGCCA 426
QY 613 GCTAACGGTACGACGCTGTAATGCTGGGGCGCTCTCTACACCGCCACTGTGAAGAAC 672
DB 427 GCTACGCGGTACGACGCTGTAATGCTGGGGGTGCTCTATGCTGCCACTGTGAAGAAC 486
QY 673 TTCCTGGGAGCTGAGCCCATCATCTCCGAGCTGTGGGTGAGCTGAGACTTGGATTCGA 732
DB 487 TACCTGGGAGCTGAGCCCATCATCTACACAGCAGTGGGTGCTGCCGAGGACTTGGATTGG 546
QY 733 ACAGAGCTTTGTCTATCTTCTGCTTAATCTCCAGCTTTGTCGACGCTATGGTCTGAGC 792
DB 547 ACAGATCTTGTCTCTCTGCTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAA 606
QY 793 CCAGCTGAGTGGGGGATGAAGATG 817
DB 607 CCAGCCGAATGGGGGATGAAGATG 631

RESULT 6

US-10-149-819-25
;; Sequence 25, Application US/10149819
;; Publication No. US20030044913A1
;; GENERAL INFORMATION:
;; APPLICANT: INCYTE GENOMICS, INC.
;; APPLICANT: YUE, Henry
;; APPLICANT: AZIMZAI, Yalda
;; APPLICANT: TANG, Y. Tom
;; APPLICANT: PATTERSON, Chandra
;; APPLICANT: BAUGHN, Mariah R.
;; APPLICANT: LU, Dyung Aina M.
;; APPLICANT: SHAH, Purvi
;; APPLICANT: LAL, Preeti
;; APPLICANT: AU-YOUNG, Janice
;; APPLICANT: BURFORD, Neil
;; TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
;; FILE REFERENCE: PF-0760 PCT
;; CURRENT APPLICATION NUMBER: US/10/149,819
;; CURRENT FILING DATE: 2002-06-10
;; PRIOR APPLICATION NUMBER: 60/172,852; 60/172,354
;; PRIOR FILING DATE: 1999-12-10; 1999-12-16
;; NUMBER OF SEQ ID NOS: 42
;; SOFTWARE: PERL Program
;; SEQ ID NO 25
;; LENGTH: 3293
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; OTHER INFORMATION: Incyte ID No. US20030044913A1 1674661CB1
US-10-149-819-25

Query Match 6.3%; Score 251.8; DB 14; Length 3293;
Best Local Similarity 52.3%; Pred. No. 8.6e-59;
Matches 706; Conservative 0; Mismatches 627; Indels 18; Gaps 6;
QY 383 GAAGCATCGACTGGATGGTACCTGAGACTTCACAGACAGAACTGCGAGGAAGAAAGCAAG- 441
DB 387 GAGCATCTCTGGGAGGCCCCCGTGGAGAAAGAACTGAGTGTATCCAGAAGGAAGA 446
QY 442 --AAGAGGAGCAATGTACAAATTTTATCCAGATTTCTGCCCATTTGTCAATGCTCTCACC 499
DB 447 ACAACGACAGCAGTGTCTCAACTTCTCGAGCCCTTCTCGAGCCCTACAACTGCTCCACC 506
QY 500 TCCTCACCTGGGCACTTCGCTTTTGTATCCAACTGCGGGTATTGTATGTGTCCAGTT 559
DB 507 TGTACGCTGTGGCACTTACGCTTCCAGCCCCAAGTGCACCTACGCTCAACATGCTCACCT 566
QY 560 TCCA---GCAGTTTGAAGACTTTGAGAGCGCGCGGGGAAATGTCTTTTTCAGCCAGCTC 616
DB 567 TCACTTTGAGCATGAGAGTTTGAAGATGGGAAGGCAAGTGTCTCTATGACCCAGCTA 626
QY 617 AACGTCAGCAGCTGTATGCTGGGGGCTCTCTACACGCCCACTGTGAAGAACTTCC 676
DB 627 AGGGCATGCTGGCCTTCTTGTGATGCTGAGTGTCTCTCGGCCACACTCAACAACACTTCC 686
QY 677 TGGGACGTCAGCCCATCATCTCCCGAGCTGTGGGTCGAGCTGAGGACTGGATTTCGAACAG 736
DB 687 TGGGACGTCAGCCCATCATCTCTGCTTAACATGGGGC---CCACCACTCCATGAAGACAG 743
QY 737 AGACCTTGTCTCTCTGGCTTAATGCTCCAGCCTTTTTCGAGCTATGCTCTGAGCCCCAG 796
DB 744 AGTACCTGGCCTTTTGGCTCAACGAACCTCACTTTGTAGGCTCTGCTCTATGATACCTGAGA 803
QY 797 CTGAGTGGGGGATGAAGATGAGACGATGAATCTTTTCTTTTCTTTCAGGAGCACTTCC 856
DB 804 CTGTGGGAGCTTTCACGGGGACGACGACAGGCTACTTCTTCTTCAGGGAGCGGGCAC 863
QY 857 GAGTGTGTGACTCTCTATGAGCGCATCAAGGTCCCAAGAGTGGCCGAGTGTGTGGGGGG 916
DB 864 TGGAGTCCGACTGCTATCCGAGCAGGAGTGGTGGCTGCTGTGCGCCCGTGTCTCGAAGGGCG 923


```
Db 1193 GCGCCCTATAGGAGTACCATGAGGAGCCAGAGTGGGACCGCTACACTGACCCT---G 1249
Qy 1217 TGCCCCAGCCAGACCTGGAGAGTGCATCGCCCAACACATGAAGCTCCAGCAGTTGGAT 1276
Db 1250 TACCCAGCCCTCGCGCTGGCTCTGCTGCTATTAACAACCTGGCATCGCGCCAGCGCTACACCA 1309
Qy 1277 CTTCACTCTCCCTGCGCAGACCGGCTGCTACACCTTTATCAGAGACCAACCTCTCATGACA 1336
Db 1310 GCTCCCTGGAGCTACCCGACAACTCTCAACTTCTGTCAGAGACACCCGCTGATGGAG 1369
Qy 1337 GCGCCCTGTTCGCGGCTGAGCGCGCCCGCTGCTGCTGCTCACTACAGATACAGCCCTATCTCA 1396
Db 1370 AGCAGTGGGCGCTCGGCTGAGCGCGCCCGCTGCTGCTGAGAGAGGCAACCACTTCACCC 1429
Qy 1397 GAGTCGTGGCCCAACAGGGTGACAGCCTCTCAGGGAAGAAATATACGTGCTCTACCTGG 1456
Db 1430 ACCTGGTGGCGGACCGGCTTACAGGACTTGATGGAGCCACTATACAGTCTGCTTCATTG 1489
Qy 1457 GGACAGAGATGACACACTCCACCGCGCTGCTGCGCATTTGGAGCTCAGCTCAGTCTTGG 1516
Db 1490 GCACAGAGACGCGCTGCTGCTCAAGGCTGTGAGCCTGGGCGCCCTGGGTTCACCTGATTG 1549
Qy 1517 AGGATCTGGCCTTGT---CCAGAAACACAGCGGCTTGAGAGCATGAAATTTGACCAAG 1573
Db 1550 AGGAGCTGAGCTGTTTGACGAGGACCCATGAGAAGCCTGGTCTATCTCAGAGCAAGA 1609
Qy 1574 ATTGGCTCTGGTGGGCTCCCACTACTAGGTGACACAAAGTGAACACCACTGTGGCC 1633
Db 1610 AGCTGCTCTTGGCGGCTCCGCTCTCAGCTGGTGCAGCTGCCGCTGGCGGCTGCATAA 1669
Qy 1634 GTCTCAGAGCTCTCGGAGTGTATCTTGGCCAGGACCGCGCTGCTGCGGCTGGAGCTTCC 1693
Db 1670 AGTATCGCTCTGTGAGAGTGTGCTCTCGCCGGGACCCCTATTGCGCTGAGGCGTCA 1729
Qy 1694 GGCTTGATGCTGTGTGGCCACGCGCGCGGA 1724
Db 1730 ACACGAGCGCTGTGTGGCGTGGGTGGCCA 1760

RESULT 8
US-10-002-050-11
; Sequence 11, Application US/10002050
; Publication No. US20030032095A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Yang, Meijia
; APPLICANT: Boldog, Ferenc
; APPLICANT: Herrmann, John
; TITLE OF INVENTION: No. US20030032095A1e1 Nucleic Acid Sequences Encoding Human Semaph
; FILE REFERENCE: 15966-554 Cura-54 CON-S14
; CURRENT APPLICATION NUMBER: US/10/002,050
; PRIOR FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 09/604,286
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,584
; PRIOR FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2155
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (166)..(1935)
US-10-002-050-11
```

Query Match 6.2%; Score 248.6; DB 14; Length 2155;
Best Local Similarity 52.1%; Pred. No. 5.4e-58;
Matches 704; Conservative 0; Mismatches 629; Indels 18; Gaps 6;

```
Qy 383 GAAGGATCGACTGGATGGTACCTGAGACTCACAGACAGAACTGCAGAGAAAGGCAAG- 441
Db 419 GAGCGATCTCTGGGAGGCGCCGCTGGAGAAGAAGACTGAGTGTATCCAGAAAGGGAAGA 478
Qy 442 --AAAGAGAGCAATGTCACAAATTTATCCAGATTTCTCGCCATTTGATCAATGCTCTCAC 499
Db 479 ACAACAGACCCGAGTCTTCAACTTCATCCGCTTCTGACGCCCTCAAAATGCTCTCCACC 538
Qy 500 TCTCTACGTCGCGCACCTTCTGCTTTTATCCGAAGTGGGGGTTATTGATGCTCTCAGTT 559
Db 539 TGTACGTCGTGGCACCTAGCCCTCCAGCCCAAGTGCACCTACGTCAACATGCTCACCT 598
Qy 560 TCCA---GCAGGTTGAAAGACTTTGAGAGCGCGCGGGGAAATGTCCTTTTGTAGCCAGCTC 616
Db 599 TCACTTTGGAGCATGAGAGTTTGAAGATGGGAAGGCAAGTCTCCCTATGACCCAGCTA 658
Qy 617 AACGGTCACAGCTGTAAATGGCTGGGGGCTCTCTACACCGCCACTGTGAAGAAGCTTCC 676
Db 659 AGGGCCATGCTGGCCTTCTTGTGGATGGTGAAGTGTACTCGGCCACACTCAACAAGCTCC 718
Qy 677 TGGGGAAGTGGCCCATCATCTCCGAGCTGTGGGTGCGAGCTGAGGACTGGATTTCGAACAG 736
Db 719 TGGGCACGAAACCCATTATCTCGCTAAACATGGGCG---CCACCACTCCATGAACACAG 775
Qy 737 AGACCTTGTCACTCTGGCTTAATGTCCAGCCCTTTGTGCGAGCTATGGTCTCGAGCCAG 796
Db 776 AGTACCTGGCTTTTGGCTCAACGAACCTCACTTTCTAGGCTCTGCTATGATACCTGAGA 835
Qy 797 CTGAGTGGGGGATGAAGATGGAGAGCATGAATCTTTTTTTTCTTACGGAGACCTCC 856
Db 836 GTGTGGGAGCTTTCAGGGGGAGCAGCAAGGTCTACTTCTTCTTTCAGGAGCGGCGAG 895
Qy 857 GAGTGTGAGCTCTTATGAGCGCATCAAGGTCCCAAGAGTGGGCCCGAGTGTGTGCGGGG 916
Db 896 TGGAGTCCACCTGCTATGCGGAGCAGTGGTGGCTGCTGCGCCGCTGCTGCAAGGGCG 955
Qy 917 ACCTTGGGGGAGGAGAACCTTTTACAGAGATGGAGCAGTGTTCGTAAGGCTGACCTGC 976
Db 956 ATATGGGGGCGCAGGACCTCTGAGTACCAGTTCGGAAGAGATCCAGCGGGTGTGAGG 1015
Qy 977 TGTGCCAGGGCGCGAGCATGCGCGGCTCTCGAGGCTATGGCAGAGCTTC 1036
Db 1016 CATGCTCTGCCCGAACTGGCAGCTTACTTCAACCACTGCGAGGCGATG---CACACCC 1072
Qy 1037 GGCCTCAGCTGGAGCGGGAACCCCATCTTTTATGGATCTTTTCTCCAGTGGGAAG 1096
Db 1073 TGCAGGACACCTCTCTGGCAACAACACACCTCTTTTGGGGTCTTCAAGCACAGTGGGG 1132
Qy 1097 GAGCTGCCATCTCTGCTGTGTGCTTCCGACCCCAAGACATCCGGGCGAGTGTGAATG 1156
Db 1133 ACATGTACCTGTGCGCCATCTGTGATACCAGTTCGGAAGAGATCCAGCGGGTGTGAGG 1192
Qy 1157 GTCCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCTGTATGGAACAGAGG 1216
Db 1193 GCGCCTATAGAGTACCATGAGGAGCCCAAGTGGGAGCGCTACACTGACCT---G 1249
Qy 1217 TGCCCCAGCCAGACTTGAGAGTGCATCGCCCAACAACTGAAGTCCAGAGTTGGAT 1276
Db 1250 TACCCAGCCCTCGGCTGCTGCTGCTGATTAACAACCTGGCATCGGCCACAGGCTACACCA 1309
Qy 1277 CCTCACTCTCCCTGCCAGACCGGCTGCTACCTTTATCAGAGACCAACCTCTCATGAGCA 1336
Db 1310 GCTCCCTGGAGCTACCCGACAACTCTCAACTTCTGTCAGGAAGACCGCTGATGGAGG 1369
Qy 1337 GCGCGGTGTTCCGCGCTGACGCGCGCCCTCTGCTGCTACTACAGATCAGAGCTATCTCA 1396
Db 1370 AGCAGTGGGCGCTCGGTGGAGCGCGCCCTCTGCTGAGGAAGGCGCAACCACTTCACCC 1429
Qy 1397 GAGTCTGGCCCAACAGGCTGACCGCTCTCAGGGAAGAAATATGACGTCTCTACCTGG 1456
Db 1430 ACCTGGTGGCGACCGGTTACAGGACTGTGAGGAGCAACCTATACAGTCTGTTCAATTG 1489
Qy 1457 GGACAGAGGATGGACACCTCCACCGGCTGTGCGCATTTGGCAGTCTGAGCTCAGTCTTGG 1516
```

```

Db 1490 GCACAGAGAGCGCTGGCTGCTCAAGGCTGTGAGCCCTGGGGCCCTGGGTTCACTGATTG 1549
Qy 1517 AGGATCTGGCTTGT---CCAGAACACAGCGGTTGAGAGCATGAATTTGACACG 1573
Db 1550 AGAGCTGACAGCTGTTGACACAGAGCCCATGAGACCTGGTGTATCTCAGAGCAAGA 1609
Qy 1574 ATTTGGCTCTGGGGCTCCATAGCTGAGTGACACAAAGTGAACACACAGCAACTGTGGCC 1633
Db 1610 AGCTGCTCTTTGGCGGCTCCGCTCTCAGCTGGTGCAGCTGCCGTGGCCGACTGCATAA 1669
Qy 1634 GTCTCCAGAGCTGCTGGAGTGTATCTTGGCCCGCAGGACCCCGTGTGGCCCTGAGCTTCC 1693
Db 1670 AGTATCGCTCTGTCAGAGCTGTCTCTCGCCGGGACCCCTATTGGCCCTGGAGCGTCA 1729
Qy 1694 GGCTTGATGCTGTGTGGCCACCGCGGCA 1724
Db 1730 ACACAGCCCTGTGTGGCCGTGGTGGCCA 1760

RESULT 9
US-10-002-304-11
; Sequence 11, Application US/10002304
; Publication No. US20030036185A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Yang, Meljia
; APPLICANT: Boldog, Ferenc
; APPLICANT: Herrmann, John
; TITLE OF INVENTION: Polynucleotides and polypeptides encoded thereby
; FILE REFERENCE: 15966-554 Cura-54 CON-S8
; CURRENT APPLICATION NUMBER: US/10/002,304
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 09/604,286
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,584
; PRIOR FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2155
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (166)..(1935)
US-10-002-304-11

Query Match 6.2%; Score 248.6; DB 14; Length 2155;
Best Local Similarity 52.1%; Pred. No. 5.4e-58;
Matches 704; Conservative 0; Mismatches 629; Indels 18; Gaps 6;

Qy 383 GAAGGATCGACTGGATGTTACCTGAGACTCACAGACAGAACTCGAGAGAAAGGCAAG- 441
Db 419 GAGCGATCTCTGGGAGGCCCCGTGGAGAAGAAGACTGAGTGTATCCAGAAAGGAAGA 478
Qy 442 --AAAGAGGAGGATGCAAAATTTATCCAGATTCTCGCCATTTGTAATGCCCTCTACC 499
Db 479 ACACACGACCGAGTGTCTCAACTTCATCGCTCTCTGACGCCCTACAATGCCCTCCACC 538
Qy 500 TCCTCAGCTGGGCACCTCTCGCTTTTATCCGAAGTGGCGGTTATTGATGTGTCACAGTT 559
Db 539 TGTACGCTGTGGCACCTACGCTTCCAGCCCAAGTGCACCTACGTCACATGCTCACT 598
Qy 560 TCCA---GCAGGTTGAAGACTTGAAGCGCGCGGGAATATGCTCTTTTGGCCAGCTC 616
Db 599 TCACCTTGGAGCTGAGAGTTTGAAGATGGAAGGCAAGTGTCCCTATGACCCAGCTA 658
Qy 617 AACGGTCAGAGCTGTAATGGCTGGGGGGTCTCTACACCGGCCACTGTGAAGAACTTCC 676
Db 659 AGGCCCATGTGGCTTCTTTGTGATGGTGTAGCTGTACTCGGGCCACACTCAACAACACTTCC 718

```

```

Qy 677 TGGGGACTGAGCCCATCATCTCCGAGCTGTGGGTGAGCTGAGGACTGATTCGAACAG 736
Db 719 TGGCAGGAAACCATATATCTCGTAACTGGGC---CCACCACTCCATGAAGACAG 775
Qy 737 AGACCTTGTCTATCTGCTTAATGCTCCAGCCTTTTGTGCGAGTATAGTCTCAGGCCAG 796
Db 776 AGTACCTTGGCTTTTGGCTCAAGCAACTCATCTTTTGTAGCTTCCCTATGTACTGAGA 835
Qy 797 CTGAGTGGGGGATGAAGATGGAGAGCATAAATCTTTTCTTCTTCCAGGAGACCTCC 856
Db 836 GTGTGGGACGCTTACGGGGGACGACAGAGGTCTACTTCTTCCAGGAGGGGCGAG 895
Qy 857 GAGTGTGGGACTCTATGAGCGCATCAAGTCCCAAGAGTGGCCGAGTGTGTGCGGGG 916
Db 896 TGGAGTCCGACTGCTATGCGGAGCAGTGTGTGCTGTGGCCCGTGTGCAAGGCG 955
Qy 917 ACCTTGGGGGAGAGACCTTTCAGCAGAGATGGAGCAGTGTCTGAAGGCTGACCTGC 976
Db 956 ATATGGGGGCGCAGGACCTTGCAGAGAAAGTGGACAGGTTCTTGAAGCGCGGTG 1015
Qy 977 TGTGCCAGGGCCGAGCATGGCGGCGCTCCGGGTTCTGCAAGGCTATGCAAGCTTC 1036
Db 1016 CATGCTTGCCTCCGAACTGGCAGCTTACTTCAACCAAGTGCAGGCGATG---CACACC 1072
Qy 1037 GGCTCAGCCTGGAGCGGAACCCCATCTTTTATGGATCTTTTCTCCAGTGGGAAG 1096
Db 1073 TGCAGGACACTCTCTGCAACACACACTTCTTTGGGGTTTTCAGCAGCAGTGGGGTG 1132
Qy 1097 GAGCTGCCATCTCTGCTGTGTGCTTCCGACCCCAAGACATCCGGGCGAGTGTGTAATG 1156
Db 1133 ACATGTACCTTGTGGCCATCTGTAGTACCAGTGTGAAGAGATGCAGCGGTGTGAGG 1192
Qy 1157 GTCCCTTAGAGAGCTAAACATGACTGCAACAGGCGACTGCCTGTCTATGCAACAGG 1216
Db 1193 GCCCTATAAGGAGTACCATGAGAGCCCAAGTGGGACCGGTACACTGACCT---G 1249
Qy 1217 TGCCCAAGCCAGACCTTGGAGAGTGCATCGCCCAACAACTGAAGCTCCAGCAGTTGGAT 1276
Db 1250 TACCCAGCCCTCGGCTGGCTCGTGCATTAACTGCGCATCGCGCCACGGCTACACCA 1309
Qy 1277 CCTCAGCTCTCCCTGCCAGACCGGCTGCTCAGCTTTATCAGAGACCCCTCTCATGACA 1336
Db 1310 GCTCCCTGGAGCTACCGCAACATCTCAACTTCGTCAGAGCAACCCCGCTGATGAGG 1369
Qy 1337 GGCCCTGTTCCCGCTGACGGCGCCCTGCTGCTACTACAGATACAGCCTATCTCA 1396
Db 1370 AGCAGTGGGGCTCGGTGAGCGCCCTGCTGCTGGAAGAGGGCAACCACTTCACCC 1429
Qy 1397 GAGTCGTGGCCACAGGCTGACAGCCTCTCAGGGAAGAATATGAGCTGCTCTACCTGG 1456
Db 1430 ACCTGTTGGCGGACCGGTTACAGGACTTGTATGAGCCACCTATACAGTGTCTTATTG 1489
Qy 1457 GGCAGAGAGTGAACACTCCACGGGCTGTGGCATTTGAGAGTACAGTCAAGTGTCTGG 1516
Db 1490 GCACAGAGAGCGGTGCTCAAGGCTGTGAGCCTTGGGCGCTTGGGTTTCACTGATG 1549
Qy 1517 AGGATCTGGCTTGT---CCAGAACACAGCGGTTGAGAGCATGAAATTTGACACG 1573
Db 1550 AGGAGCTGACAGCTTTTGACACAGAGCCCATGAGAGCCCTGGTGTCTATCTCAGAGCAAGA 1609
Qy 1574 ATTGGCTCTCTGGTGGCTCCCATCTACTGAGGTGACACAAAGTGAACACCACTGTGGCC 1633
Db 1610 AGCTGCTCTTTGGCGGCTCCGCTCTCAGCTGGTGTGAGCTGCCGTGGCGGAGTGCATAA 1669
Qy 1634 GTCTCCAGAGCTGCTGGAGTGTATCTTGGCCCGCAGGACCCCGTGTGGCCCTGAGCTTCC 1693
Db 1670 AGTATCGCTCTGTCAGAGCTGTCTCTCGCCGGGACCCCTATTGCGCTTGGAGCGTCA 1729
Qy 1694 GGCTTGATGCTGTGTGGCCACCGCGGCA 1724
Db 1730 ACACAGCCCTGTGTGGCGGTGGTGGCCA 1760

```

RESULT 10

US-10-003-152-21
; Sequence 21, Application US/10003152
; Publication No. US20020151494A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Yang, Meijia
; APPLICANT: Boldog, Ferenc
; APPLICANT: Herrmann, John
; TITLE OF INVENTION: No. US20020151494A1el Amino Acid Sequences for Human Semaphorin-1
; FILE REFERENCE: 15966-554 Cura-54 CON-S12
; CURRENT APPLICATION NUMBER: US/10/003.152
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 09/604, 286
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140, 584
; PRIOR FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 2156
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (166)..(2037)
; NAME/KEY: Variation
; LOCATION: (1)..(2156)
; OTHER INFORMATION: N may be any nucleotide
US-10-003-152-21

Query Match 6.2%; Score 248.6; DB 13; Length 2156;
Best Local Similarity 52.1%; Pred. No. 5.4e-58;
Matches 704; Conservative 0; Mismatches 629; Indels 18; Gaps 6;

QY	383	GAAGGATCGACTGGATGGTACCTGAGACTCAGACAGAACTCAGAGGAAGAAAGCAAG- 441
DB	419	GAGCGATCTCTGGGAGCCCCCGTGGAGAGAGACTGAGTGATCCAGAAGGGA 478
QY	442	--AAGAGGAGATGTCACAAATTTATCCAGATTCCTCGCCATGTCAATGCCTCTACC 499
DB	479	ACAACGACGAGGTGCTTCAACTTCATCCGCTTCTCGAGCCCTACAAATGCCTCCACC 538
QY	500	TCCTCAGCTGGGCACTTCGCTTTTGATCCGAAGTGGGGTATTGATGTGCCAGTT 559
DB	539	TGTAGCTGTGGGCACCTTACGCCCTTCCAGCCCAAGTGCACCTACGTCACATGCTCACC 598
QY	560	TCCA---GCAGGTTGAAGACTTGAGAGCGGCGGGGAAATGTCTTTTGAGCGCAGTCC 616
DB	599	TCACCTTGGAGCATGGAGAGTTTGAAGATGGGAAGGCAAGTGTCCCTATGACCCAGCTA 658
QY	617	AACGGTCAAGCAGCTGTAATGGCTGGGGGGCTCTCTACACCGCACTGTGAAGAACTTCC 676
DB	659	AGGGCCATGCTGGCCTTCTTGGATGGTGAAGTGTACTCGGCCACACTCAACAACCTTCC 718
QY	677	TGGGAGCTGAGCCCATCATCTCCGAGCTGTGGGTGAGCTGAGGACTGGATTGCAACAG 736
DB	719	TGGGACCGAAGACCATTTATCTGCTTAACATGAGGGG---CCACACACTCCATGAAGACAG 775
QY	737	AGACCTTGTATCTCTGGCTTAATGCTCCAGCCTTTGTGCGAGCTATGCTCTGAGCCACAG 796
DB	776	AGTACCTGGCCTTTTGGCTCAACGAACTCACTTTGTAGGCTCTGCCTATGTACTCTGAGA 835
QY	797	CTGAGTGGGGGATGAAGATGGAGACGATGAATAATCTTTTCTTCTCACGGAGACCTTCC 856
DB	836	GTGTGGGAGCTTCAACGGGGACGACGACAAAGGTCTACTTCTTCTCAGGGAGGGGGCAG 895
QY	857	GAGTGTGTGAGCTCTATGAGCGCATCAGGTGCCAAGAGTGGCCGAGTGTGCGGGGG 916
DB	896	TGGAGTCGCACTGTATGCCAGCAGGAGTGGTGGCTCGTGTGGCCCGCTGTCTGCAAGGGCG 955

QY	917	ACCTTTGGGGCAGGAAGACCTTTCAGCAGAGATGACGACGTTTCTGAAGGCTGACCTGC 976
DB	956	ATATGGGGGGCGCAGCGACCTTGCAGAGGAAGTGACACAGTTCCTGAAGCGCGCGCTGG 1015
QY	977	TGTGCCCAGGGCCGAGCATGGCCGGGCTCTCGGGGTTCTCGAGGCTATGCGAGAGCTTC 1036
DB	1016	CATGCTCTGCCCCGAACTTGGCAGCTCTACTTCAACAGCTGCGAGCGATG---CACACCC 1072
QY	1037	GGCCTCAGCCTGGAGCGGGAACCCCATCTTTTATGGGATCTTTTCCCTCCAGTGGGAAG 1096
DB	1073	TGCAGGACACCTCTCTGTCGCTTCCGACACACCTCTTTTGGGGTTTTTCAACAGAGTGGGGTG 1132
QY	1097	GAGCTGCCATCTCTGCTGTGTGCTTCCGACCCCAAGACATCCGGGACGTGCTGAATG 1156
DB	1133	ACATGTACCTCTGCGCCATCTGTGAGTACCAGTTTGAAGAGATCCAGCGGTTGTTGAGG 1192
QY	1157	GTCCCTTTAGAGAGCTAAACATGACTGCACAGGGGACTGCTGTCTATGACACACAGAGG 1216
DB	1193	GCCCTATTAAGAGGTACCATGAGGAAGCCAGAGTGGGAGCGCTTACACTGACCT---G 1249
QY	1217	TGCCCCAGCCAGACCTGAGAGTGCATCGCCCAACATGAAGCTCCAGCAGTTTGGAT 1276
DB	1250	TACCCAGCCCTCGGCTGGCTGCTGCAATTAACAACCTGGCAGCGGCCACGGCTACACCA 1309
QY	1277	CCTCACTCTCCCTGCCAGACCGCGTCTCACTTTTATCAGAGACACACCTCTCTCATGGACA 1336
DB	1310	GCTCCCTGGAGTACCCGACAACTCCTCAACTTCGTCAGNAGACACCCGCTGATGGAGG 1369
QY	1337	GGCCGCTGTCCCGGCTGACGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1396
DB	1370	AGCAGGTGGGCTCGGTGGAGCGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1429
QY	1397	GAGTCGTGGCCACAGGCTGACAGCCCTCTCAGGGAAGAAATATGACCTGCTCTACCTGG 1456
DB	1430	ACCTGGTGGCGACCGGCTTACAGACTTGATGAGGCCACCTATACAGTGTCTGTTCAATTG 1489
QY	1457	GGACAGAGATGGACACCTCCACCGGCTGTGCGCATTTGGAGCTCAGCTCAGTGTCTTGG 1516
DB	1490	GCACAGGAGACGGCTGGCTGCTCAAGGCTGTGAGCCTGGGCGCTTCACTGATG 1549
QY	1517	AGGATCGCCCTGTT---CCAGAACACAGCGCGTTGAGAGCATGAATAATGTACACAG 1573
DB	1550	AGGAGCTGCAGCTGTGTTGACCAAGAGGCCCATGAGAAGCTGTGCTATCTCAGAGCAAAA 1609
QY	1574	ATTGGCTCTGTGGCTCCCATCTAGAGTGACAAAGTGAACACCAGCAACTGTGCCC 1633
DB	1610	AGCTGCTCTTTGGCGGCTCCCGCTCTCAGCTGGTGCAGCTGCCGTGGCCGACTGCATTA 1669
QY	1634	GTCTCCAGAGCTGCTCGGAGTGTATCTTGGCCCAAGGACCCGCTGTGCGCTGGAGCTTCC 1693
DB	1670	AGTATCGCTCTGTGCAGACTGTGTCTCTCGCCCGGACCCCTATTTCGCCCTGGAGCGTCA 1729
QY	1694	GGCTTGATGCTGTGTGGCCACCGCCGCGA 1724
DB	1730	ACACCAGCGCTGTGTGGCCGTGGTGCGCA 1760

RESULT 11

US-10-002-050-21
; Sequence 21, Application US/10002050
; Publication No. US20030032095A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Yang, Meijia
; APPLICANT: Boldog, Ferenc
; APPLICANT: Herrmann, John
; TITLE OF INVENTION: No. US20030032095A1el Nucleic Acid Sequences Encoding Human Se
; FILE REFERENCE: 15966-554 Cura-54 CON-S14
; CURRENT APPLICATION NUMBER: US/10/002.050
; CURRENT FILING DATE: 2001-11-02

US-10-002-304-21

Query Match	6.2%	Score 248.6	DB 14	Length 2156
Best local Similarity	52.1%	Pred. No. 5.4e-59		
Matches	704	Conservative	0	Mismatches 629; Indels 18; Gaps 6;
QY	383	GAAGGATCGACTGATGGTACCTGAGACTCACAGACAGAAGCTGCAGGAAGAAAGGCAAG- 441		
DB	419	GAGGATCTCTGGGAGGCCCGGTGGAGAAGAAAGACTGAGTGTATCCGAAGAGGGAAGA 478		
QY	442	--AAAGAGACGAATGTACAATTTTATCCAGATTTCTGCCAATGTGTCAATGCCCTCTCACC 499		
DB	479	ACAACCAGACCGAGTGTCTCAACTTCTCCGCTTCTGCGACCTCAATAGCTCTCCCAACC 538		
QY	500	TCCTCAGTCGGGCACCTTCGCTTTTCATCCGAAGTCGGGGTATTGATGTGTCCAGTT 559		
DB	539	TGTACGTCTGTGGCACTACGCTTCCAGGCCCAAGTCACCTACGTCACATGCTCACT 598		
QY	560	TCCA--CGAGGTTGAAAGACTTGCAGAGCGCCGGGAAATGCTCTTTTGAGCGACGCTC 616		
DB	599	TCACTTTGGAGCATGGAGAGTTTGAAGATGGGAAGGCAAGTGTCCCTATGACCCAGCTA 658		
QY	617	AAGGTCAGACGCTGTAATGGCTGGGGGCTCTCTACACGGCAGCTGTGAAGACTTCC 676		
DB	659	AGGGCCATGCTGGCCTTCTTGTGGATGGTGAAGTGTACTCGGCACACTCAACAACCTTCC 718		
QY	677	TGGGACTGAGCCATCATCTCCCGAGCTGTGGTGCAGCTGAGGACTGATTCGACAC 736		
DB	719	TGGCACGGAACCATATCTCGGTACATGGGGC---CCACCACTCCATGAAGACAG 775		
QY	737	AGACCTTGTCTGCTGCTTAATGCTCCAGCCTTTGTGCGAGCTATGGTCTCAGCCACG 796		
DB	776	AGTACCTGGCCTTTTGGCTCAACGAACCTCACTTTGTAGGCTCTGCCTATGTACCTGAGA 835		
QY	797	CTGAGTGGGGGATGAAGATGGAGACGATCAAAATCTTTTTTTTCTTCCACGGAGACTCCC 856		
DB	836	GTGTGGGCAGCTTCAAGGGGAGACAGCAAGGTCTACTTCTTCTCAGGAGCGGGCAG 895		
QY	857	GAGTGTGTGGACTCCTATGAGCGCATCAAGTCCCAAGAGTGGGCCGAGTGTGCGGGGG 916		
DB	896	TGGAGTCCGACTGCTATGCCGAGCAGGTGCTGCTGTGGCCGTGTCTGCAAGGCGC 955		
QY	917	ACCTTGGGGCAGAGACACCTTTCAGCAGAGATGGAGCAGTTCGTGAAGGCTGACCTGC 976		
DB	956	ATATGGGGGCGCACCGACCTCAGAGGAAGTGGACAGTTTCTTGAAGCGCGCGCTGG 1015		
QY	977	TGTGCCCAGGCGCCGAGCATGGCGGGCTTCTCGAGGCTATGGCAGAGCTTC 1036		
DB	1016	CATGCTCTGCCCGAACTGCGAGCTCTACTTCAACCAAGCTGCGAGGCGATG---CACACC 1072		
QY	1037	GGCCTCAGCCTGGAGGGGGAACCCCACTTTTATGGGATCTTTTCTCCAGTGGGAAG 1096		
DB	1073	TGCAGGACACCTCCTGSCACAACACCACTTCTTTGGGGTTTTTCAAGCACAGTGGGGTG 1132		
QY	1097	GAGTCCCATCTCTGCTGTGTGTCCTTCGACCCCAAGACATCCGGGCAGTCTGATG 1156		
DB	1133	ACATGTACCTGTCTGGGCATCTGTGAGTACCAGTTGGAAGAGATCCACGCGGTGTTTGAGG 1192		
QY	1157	GTCCCTTTAGAGAGCTAAAACATGACTGCAACAGGGGACTTCGCTGTCTATGGCAACAGG 1216		
DB	1193	GCCTCTATAGGAGTACCATGAGGAAGCCAGAAAGTGGGACCGCTACACTGACCT---G 1249		
QY	1217	TGCCCAAGCCAGACCTGGAGAGTGTGATCCCAACAAACATGAAGCTCCACGAGTTTGGAT 1276		
DB	1250	TACCCAGCCCTCGCCTGGCTCGTGTGATTAACAACCTGGCATCGCGGCCAGCGGTACACCA 1309		
QY	1277	CCTCACTCTCCCTGCCAGACCGGTGCTCACTTTTATCAGAGACCAACCCCTCTCATGGACA 1336		
DB	1310	GCTCCCTGGAGCTACCCGACAAATCCTCAACTTGTCTCAGAGACACCCGCTGATGGAGG 1369		
QY	1337	GGCCCGTGTTCGGGCTGACGGCCGCCCTGCTGTGCTACTACAGATACAGCCTATCTCA 1396		
DB	1370	AGCAGTGGGGCCTCGGCTGGAGCGGCCCTGCTGCTGTGAAGAGGGCACCACCTTCACC 1429		

Qy	1397	GAGTCGTGGCCACAGGGTACACAGCCTCTCAGGGAAGAAATATGACGTGCTCTACCTGG	1455
Db	1430	ACCTGTGTGGCCGACCGGGTTACAGGACTTGTAGGAGCCACCTATACAGTGTCTTTCATTG	1489
Qy	1457	GGACAGAGGATGACACCTCCACGGGCTGTGGCGATTGGAGCTCAGCTCAGTGTCTTGG	1516
Db	1490	GCACAGAGAGGCTGGCTGCTCAAGGCTGTGAGCCTGGGGCCCTGGGTTCACTCATTTG	1549
Qy	1517	AGGATCTGGCTTGT---CCAGAAACACAGCCGGTTGAGAGCATGAAATTGTACCAACG	1573
Db	1550	AGGAGCTGACGTGTTTGACAGGAGGCCATGAGAAGCCTGGTGCTATCTCAGAGCAAAA	1609
Qy	1574	ATTGGCTTCGTGGGCTCCCATCTAGAGTGTGACACAAAGTGAACACCACTGTGGCC	1633
Db	1610	AGCTGTCTTTTGGCGGCTCCCGTCTCAGCTGTGTGAGCTGCCCGTGGCCGACTGCATTA	1669
Qy	1634	GTCTCCAGAGCTGTGTGGAGTGTATCTGTGCCAGGACCCCGTGTGGCTGGAGCTTCC	1693
Db	1670	AGTATCGCTCTGTGACAGCTGTGTCTGTGCCCGGAGACCCCTATTGGGCTGGAGGCTCA	1729
Qy	1694	GGCTTGATGTTGTGTGGCCACGCCGGCA	1724
Db	1730	ACACGACCGCTGTGTGGCGGTGGGTGGCCA	1760
RESULT 13			
US-10-003-152-13			
; Sequence 13, Application US/10003152			
; Publication No. US20020151494A1			
; GENERAL INFORMATION:			
; APPLICANT: Shinkets, Richard			
; APPLICANT: Fernandes, Elma			
; APPLICANT: Vernet, Corine			
; APPLICANT: Yang, Meijia			
; APPLICANT: Boldog, Ferenc			
; APPLICANT: Hermann, John			
; TITLE OF INVENTION: No. US20020151494A1el Amino Acid Sequences for Human Sem			
; FILE REFERENCE: 15966-554 Cura-54 CON-S12			
; CURRENT APPLICATION NUMBER: US/10/003,152			
; CURRENT FILING DATE: 2001-11-02			
; PRIOR APPLICATION NUMBER: 09/604,286			
; PRIOR FILING DATE: 2000-06-22			
; PRIOR APPLICATION NUMBER: 60/140,584			
; PRIOR FILING DATE: 1999-06-23			
; NUMBER OF SEQ ID NOS: 49			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 13			
; LENGTH: 2284			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (166)..(1953)			
; NAME/KEY: variation			
; LOCATION: (1)..(2284)			
; OTHER INFORMATION: N may be any nucleotide			
US-10-003-152-13			
Query Match 6.2%; Score 248.6; DB 13; Length 2284;			
Best Local Similarity 52.1%; Pred. No. 5.5e-58;			
Matches 704; Conservative 0; Mismatches 629; Indels 18; Gaps 6;			
Qy	383	GAGGATFCGACTGGATGCTACCTGAGACTCAGACAGAACTCCAGAGAAAGGCAAG-	441
Db	419	GAGCGATTCCTTGGGAGGCCCGCTGGAGAAAGAGACTGAGTGTATCCGAAAGGGAAGA	478
Qy	442	--AAAGAGGACGAATGTCACAAATTTATCCAGATTCTCGCCATTGTCAATGGCTCTCAC	499
Db	479	ACRACAGACCGAGTGTCTCACTTCTATCCGCTTCTGAGGCCCTACAAATGCCCTCCACC	538
Qy	500	TCCTCAGCTGCGGCACCTTCGCTTTTGTATCCGAAGTCCGGGGTTATTGTATGTGTCACGTT	559

Db 539 TGTACGTCGTGGCAACCTACGCCCTTCCAGCCCAAGTGCACCTACGTCACAACTGCTCACT 598
Qy 560 TCCA---GCAGGTTGAAGACTTGTAGAGCGCCGGGGAATATGCTTTTTCAGCCAGCTC 616
Db 599 TCACCTTTGGACATGGAGAGTTTGAAGATGGGAAGGCAAGTGTCCCTATGACCACGCTA 658
Qy 617 AACGGTCAGCAGCTGTATATGCTGGGGGCTCTCTACACCGCCACTGTGAAGAACTTCC 676
Db 659 AGGGCCATGCTGGCCCTCTCTGTGGATGGTGAAGTGTACTCGGCCACACCTCAACAACCTCC 718
Qy 677 TGGGAGTACGCCCATCATCTCCGAGCTGTGGTTCGAGCTGAGGACTGGATTGCAACAG 736
Db 719 TGGGACGGAACCCATTATCTCGTAACTTACATGGGGC---CCACCACTCTCAATGAAGAC 775
Qy 737 AGACCTTGTCTATCTGGCTTAATGCTCCAGCCTTTGTGCGAGCTATGCTGCTGAGCCCA 796
Db 776 AGTACCTGGCCCTTTTGGCTCAACGNACTCACTTTGTAGGCTCTGCCCTATCTACCTGAGA 835
Qy 797 CTGAGTGGGGGATGAAGATGGAGCAGTAAATCTTTTTCACGAGGACTGACCTGCC 856
Db 836 GTGTGGGAGCTTTCACGGGGACGACGAAAGGCTACTTCTTTCAGGAGCGGGCAG 895
Qy 857 GAGTGTGGACTCTATGAGCGCATCAAGTCCCAAGAGTGCCTCGAGTGTGTCGGGGG 916
Db 896 TGGAGTCGCACTGTATGCCGAGCAGTGTGGTCTGTGTCGGCCGTCTGCAAGGGCG 955
Qy 917 ACCTTGGGGGAGGAACCCCTTCAGCAGATGGAGCAGCTTTCTGAAGCTGACCTGCG 976
Db 956 ATATGGGGGCGCAGGACCTGTCAGAGAGTGGACCACTGCTGTAAGCGGGCTGG 1015
Qy 977 TGTGCCAGGGCCGAGCATGCGCGGCTCCGGGTTCTCAGGCTATGCGAGAGCTTC 1036
Db 1016 CATGCTGCGCCGAACCTGGCAGCTCTACTTCAACCACTGCGCGATG---CACACCC 1072
Qy 1037 GGCCTCAGCTGGAGGGGAACCCCATCTTTTATGGATCTTTTCTCCAGTGGGAAG 1096
Db 1073 TGCAGGACACCTCTCTGSCACACACACCTTTCTTTGGGGTTTTCAGACAGCTGGGTG 1132
Qy 1097 GAGCTGCCATCTCTGCTGTGTGCTTCCGACCCCAAGACATCCGGCAGCTGCTGAATG 1156
Db 1133 ACATGTACCTGTGGCCATCTGTGATACCACTTGGAGAGATCCAGCGGTGTTGAGG 1192
Qy 1157 GTCCCTTTAGAGATAAACATGACTGCAACAGGGGACTCGCTCTCATGGAACACGAGG 1216
Db 1193 GCCCTATAAGGAGTACCATGAGGAAGCCAGAGTGGGACCGCTACACTGACCT---G 1249
Qy 1217 TGCCCCAGCCAGCTGGAGTGCATCGCCCAACACATGAAGTCCAGCAGTTGGAT 1276
Db 1250 TACCCAGCCCTCGGCCCTGGCTCGTGCATTAACAACCTGGCATCGGCCCAACGCTACACCA 1309
Qy 1277 CCTCACTCTCCCTGCGACACCGGTGCTCACTTTATCAGAGACCACTCTCATGGACA 1336
Db 1310 GCTCCTGGAGCTACCCGACAACTCTCACTTGTGTCAGAAAGCACCCTGATGGAGG 1369
Qy 1337 GGGCCGTGTCGCCGTGACGCCCGCCCTGCTGGTCACTACAGATACAGCCCTATCTCA 1396
Db 1370 AGCAGTGGGGCCCTCGGTGGAGCGCCCTGCTGCTGAAGAAGGGCACTTCAACC 1429
Qy 1397 GAGTCGTGGCCCAAGGCTGACAGCTCTCAGGGAAGAATATGACGTCTACCTGG 1456
Db 1430 ACCTGGTGGCCGACCGGTTACAGGACTTGTAGGACCACTATACAGTGTGTTCAATTG 1489
Qy 1457 GGACAGAGATGGACACTCCACCGGCTGTGCCATTTGAGCTGAGCTCAGCTAGTGTCTGG 1516
Db 1490 GCACAGGAGAGCGGTGGCTGTCAAGGCTGTGAGCTGGGGCCCTGGGTTCACTGATG 1549
Qy 1517 AGGATCTGGCCCTTGT---CCAGAAACCAAGCCGGTTGAGAGCATGAATTTACCAACG 1573
Db 1550 AGGAGCTGCACTGTTTGAACGAGGAGCCATGAGAGCCTGGTCTATCTCAGAGCAAGA 1609
Qy 1574 ATTGGCTCCTGGTGGGCTCCCATATCTAGGTGACACAAAGTGAACACCAAGCACTGTGCC 1633
Db 1610 AGCTGCTCTTGGCGGCTCCCGCTCTCAGCTGGTGCAGCTGCCCGTGGCCGACTGCATAA 1669

RESULT 14

US-10-002-050-13
; Sequence 13, Application US/10002050
; Publication No. US20030032095A1

GENERAL INFORMATION:

; APPLICANT: Shinkets, Richard
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Yang, Melija
; APPLICANT: Boldog, Ferenc
; APPLICANT: Herrmann, John
; TITLE OF INVENTION: No. US20030032095A1el Nucleic Acid Sequences Encoding Human Se
; FILE REFERENCE: 15966-554 Cura-54 CON-514
; CURRENT APPLICATION NUMBER: US/10/002,050
; CURRENT FILING DATE: 2001-11-02
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,584
; PRIOR FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 13

; LENGTH: 2284

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (166)..(1953)

; NAME/KEY: variation

; LOCATION: (1)..(2284)

; OTHER INFORMATION: N may be any nucleotide

US-10-002-050-13

Query Match

Best Local Similarity 6.2%; Score 248.6; DB 14; Length 2284;

Matches 704; Conservative 0; Mismatches 629; Indels 18; Gaps 6;

Qy 383 GAAGATCGACTGGATGTACTGAGACTCACAGACTGACAGAGAACTGCGAGGAAGGCAAG- 441
Db 419 GAGCGATCTCCTGGGAGGCCCGCTGGAGAAGAGACTGAGTGTATCCAGAAGGGAAGA 478
Qy 442 --AAGAGGACGAATGTCACAAATTTATCCAGATTCTCGCATTTCTCAATGCCTCTCACC 499
Db 479 ACAACCAAGCAGAGTGTCTCAACTTCATCCGCTTCTCGAGCCCTACATGCTCCACC 538
Qy 500 TCCTCACCTCGGGACCTTCGCTTTTGATCCGAAGTCGGGTATTGATGTGTCCAGTT 559
Db 539 TGTAGCTGTGGCCACTACGCTTCCAGCCCAAGTGCACCTACATGCTCACTC 598
Qy 560 TCCA---CGAGTTGAAAGACTTGAGAGCGCCGGGGAATGTCTTTTGTGACGAGCTC 616
Db 599 TCACCTTGGAGCATGGAGAGTTTGAAGATGGGAAGGCAAGTGTCTCTATGACCCAGCTA 658
Qy 617 AACGTCAGCAGCTGTAATGGCTGGGGGCTCTCTACACCGCCACTGTGAAGAACTTCC 676
Db 659 AGGGCCATGCTGGCCCTCTTGTGGATGGTGAAGTGTCTACTCGGCCACACTCAACAAC 718
Qy 677 TGGGACTGAGCCCATCATCTCCGAGCTGTGGGTGCGAGCTGAGGACTGGATTCCGAACAG 736
Db 719 TGGGACCGAAGACCCATTATCTGCTGAACATGGGGC---CCACCACTCTCATGAAGACAG 775
Qy 737 AGACCTTGTCTCTGCTTAAATGCTCCAGGCTTTGTCGCACTATGCTCTGAGCCAG 796


```
Db 1016 CATGCTCTGCCCGCAACTGGAGCTCTACTTCAACAGCTGCAGGCGCATG---CACACCC 1072
Qy 1037 GGCCTCAGCTGGAGCGGGAACCCCATCTTTTATGGATCTTTTCTCCAGTGGGAAG 1096
Db 1073 TGCAGGACACTCTCTGCGACAACACCACTTCTTTGGGGTTTTCAGACACAGTGGGGTG 1132
Qy 1097 GAGCTGCCATCTCTGCTGTGTGTGCTTCCGACCCCAAGACATCCGGGACAGTGTGAATG 1156
Db 1133 ACATGTACCTCTGCTGGCCATCTGTGAGTACCAGTTGGAAGAGATCCAGCGGTGTTGAGG 1192
Qy 1157 GTCCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCCCTGTCTATGGACAACGAGG 1216
Db 1193 GCCCTATAGAGGTACCATGAGGAAGCCAGAAAGTGGGACCGCTACACTGACCCCT---G 1249
Qy 1217 TGCCCCAGCCAGCACTGGAGAGTGCATCGCCAAACAATGAAGCTCCAGCAGTTTGGAT 1276
Db 1250 TACCCAGCCCTCGGCCCTGGCTCGTGCATTAACTACTGGCATCGGCGCCACGGCTACACCA 1309
Qy 1277 CCTCAGCTCTCCCTGCCAGACGGGTGCTCAGCTTTATCAGAGACCACTCTCTCATGGACA 1336
Db 1310 GCTCCCTGGAGCTACCCGACAACATCTCAACTTCGTCAAGAAGCACCCGCTGATGGAGG 1369
Qy 1337 GGCCCGTGTCCCGGCTGACGGCGGCCCTGCTGTGCTCACTACAGATACAGCCTATCTCA 1396
Db 1370 AGCAGTGGGGCCTCGGTGAGCGGCCCTGCTGCTGGAAGAGGGCAACCACTTCACCC 1429
Qy 1397 GAGTCTGGGCCACAGGGTGACAGCCTCTCAGGGAAGAATATGAGCTGCTCTACCTGG 1456
Db 1430 ACCTGGTGGCGGACCGGTTACAGGACTTGATGGAGCCACCTATACAGTCTGCTTCATTG 1489
Qy 1457 GGACAGAGGATGGACACTCCACCGGGCTGTGCGCATTTGGGAGCTCAGTCTCATTGG 1516
Db 1490 GCACAGGAGCGGCTGGCTGCTCAAGGCTGTGAGCCTGGGGCCCTGGGTTACCTGATTG 1549
Qy 1517 AGGATCTGGCCTTGT---CCCAAGAACACAGCCGGTTGAGAGCATGAATTTGACCAG 1573
Db 1550 AGGAGCTGCAGCTGTTTGACAGGAGGCCCATGAGAAGCCTGGTCTATCTCAGAGCAAGA 1609
Qy 1574 ATTGGCTCCTGGTGGGCTCCCATACTGAGGTGACACAAGTGAACACCAGCAACTGTGGCC 1633
Db 1610 AGCTGCTCTTTGCCGGCTCCCGCTCTCAGCTGGTGCAGCTGCCCGTGGCCGACTGCATAA 1669
Qy 1634 GTCTCCAGAGCTGCTGGAGTGTATCCTGGCCCAAGACCCGGTGTGCGCCTGGAGCTTCC 1693
Db 1670 AGTATCGCTCCTGTGCAGACTGTGCTCCTCGCCGGGACCCCTATTGGCCCTGGAGCGTCA 1729
Qy 1694 GGCTTGATGCTGTGTGCGCCACCGCGCGA 1724
Db 1730 ACACAGCGCTGTGTGGCGGTGGGCCA 1760
```

Search completed: September 26, 2003, 11:18:05
Job time : 895.613 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2003, 19:14:04 ; Search time 8343.28 Seconds
(without alignments)
11429.597 Million cell updates/sec

Title: US-09-284-180A-2

Perfect score: 2331

Sequence: 1 atgtctgccaggccgagcg.....gtgatgagacctcatctaa 2331

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_on.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vit.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2331	100.0	4008	10	AB002563	AB002563 Rattus no
2	2150	92.2	4002	10	AB021291	AB021291 Mus muscu
3	2004	86.0	2354	10	AF038652	AF038652 Mus muscu
4	1827	78.4	2948	9	BC038411	BC038411 Homo sapi
5	1827	78.4	2971	6	AX136143	AX136143 Sequence
6	1827	78.4	2971	6	BD123524	BD123524 Secretary
7	1827	78.4	2971	9	AK075384	AK075384 Homo sapi
8	1827	78.4	4293	9	HS0801526	AL136552 Homo sapi
9	1626	69.8	2893	6	AX003081	AX003081 Sequence
10	1626	69.8	2893	6	BD177663	BD177663 SBSEMN1 p
11	1626	69.8	2893	6	E35443	E35443 SBSEMN1 pol
12	1626	69.8	2894	6	AX003083	AX003083 Sequence
13	1626	69.8	2894	6	BD177664	BD177664 SBSEMN1 p
14	1626	69.8	2894	6	E35444	E35444 SBSEMN1 pol
15	1441	61.8	2672	9	AB021292	AB021292 Homo sapi
16	1190	51.1	2507	9	BC018361	BC018361 Homo sapi
17	1188	51.0	2400	9	AB022317	AB022317 Homo sapi
18	1023	43.9	1913	9	AF053369	AF053369 Homo sapi
19	654	28.1	799	6	AX003085	AX003085 Sequence
20	654	28.1	799	6	BD177665	BD177665 SBSEMN1 p
21	654	28.1	799	6	E35445	E35445 SBSEMN1 pol
22	616	26.5	2669	9	AB047604	AB047604 Macaca fa
23	569	24.4	250709	2	AC135520	AC135520 Rattus no
24	569	24.4	3205	10	AB022311S6	AB022316 Mus muscu
25	569	24.4	109332	2	AC134899	AC134899 Mus muscu
26	569	24.4	162691	10	AC003061	AC003061 Mouse Chr
27	569	24.4	214837	10	AC007305	AC007305 Mus muscu
28	466	20.0	64154	2	AL451003_3	Continuation (4 of
29	466	20.0	110000	2	AL451003_2	Continuation (3 of
30	466	20.0	126228	9	AC006543	AC006543 Homo sapi
31	466	20.0	234523	9	AC006544	AC006544 Homo sapi
32	454	19.5	163903	9	AC007387	AC007387 Homo sapi
33	372	16.0	669	6	AX136484	AX136484 Sequence
34	372	16.0	669	6	BD123724	BD123724 Secretary
35	296	12.7	2406	9	AK095587	AK095587 Homo sapi
36	251	10.8	3293	6	AX174731	AX174731 Sequence
37	248	10.7	2155	6	AX060303	AX060303 Sequence
38	248	10.7	2156	6	AX060313	AX060313 Sequence
39	248	10.7	2284	6	AX060305	AX060305 Sequence
40	248	10.7	2646	6	AX0704763	AX0704763 Sequence
41	248	10.7	3776	6	BD171174	BD171174 Novel gen
42	248	10.7	3776	9	AB051526	AB051526 Homo sapi
43	247	10.6	2558	6	AX528277	AX528277 Sequence
44	247	10.6	3112	6	AX528279	AX528279 Sequence
45	247	10.6	3781	6	AX376386	AX376386 Sequence

ALIGNMENTS

RESULT 1

AB002563

LOCUS

DEFINITION

AB002563

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

1 (sites)

REFERENCE

AUTHORS

AB002563 Rattus norvegicus mRNA for semaphorin W, complete cds.
4008 bp mRNA linear ROD 09-APR-1999

AB002563

AB002563.1 GI:4519426

semaphorin W.

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

Encinas, J.A., Kikuchi, K., Chedotal, A., de Castro, F., Goodman, C.S.

and Kimura, T.

TITLE Cloning, expression, and genetic mapping of Sema W, a member of the semaphorin family
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (5), 2491-2496 (1999)
MEDLINE 99162633
PUBMED 10051670
REFERENCE 2 (bases 1 to 4008)
AUTHORS Kimura,T. and Kikuchi,K.
TITLE Direct Submission
JOURNAL Submitted (31-MAR-1997) Toru Kimura, Sumitomo Pharmaceuticals Research Center, Discovery Research Laboratories II; 1-98, Kasugade Naka 3-chome, Konohana-Ku, Osaka, Osaka 554, Japan.
(E-mail:tkimura@sumitomopharm.co.jp, Tel:06-466-5228, Fax:06-466-5491)

FEATURES

Source

Location/Qualifiers
1..4008
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/sex="Female"
/tissue_type="Brain"
/dev_stage="Adult"
76..2406
/codon_start=1
/product="semaphorin W"
/protein_id="BAA75629.1"
/db_xref="GI:4519427"

CDS

/translation="MLARAERPRGPRPPVPPPLSLLLALLSAPVCGVRPRS
VPTSLPISEADSYLTREASHYNYKSYALLVDTPASHLYVNGRDSIFALTLPFSGERP
RRLDMVETHRONKRGKKKEDECHNFILAIYNASHLLTGCDFDPKGVLDVS
SPQVERLESNGKCPFEPAQSAAMAGVLYATVKNFLGTEPIISRAVGRADWDI
RPTETLSMLNAPAFVAAWVLSPEAWGDDGDEIFFTETSRVLDSTRIKVPVRAR
VCAGDLGRKTLQQRWTFLLKADLLCPGEHGRASGVQLQAMAEPLRQPGAGTPIFYGI
FSSOEGAAISAVCAFRPQDIRAVLNGPFEHLKHDNRLGRLPVMDEVPQPRGECIAN
NMKLOQSSLSLDRVLTFIRDPLMDRPVFPADGRLLVTDYALRVVAHRVTSL
SKKEYDVLVLTGDEGHLRAVRIGALSVLEDLALFPEPQPVESMKLHDMLLVGSHT
EVQYNTNLCRGSCSECLIAQDPVCWSFRLDACVAHAGEHRHMGQDIESADVSSL
CPKEPGEHPVFEVPAIVGVHVLPCSPSSAWASCVHQPQSGVTALTAPRRDGLLEVYVT
PGMAGYACEQEGGAARVAAVSLVWSQGRPSNRATVVGAGLVGFLGLVLAASLT
LLLIQRQRRLRLARDKVGLDLGAPPSTTSYSDPPSPSPEDERLPLALKRG
SGFGFPPFLDSCPSAHLRLTGLAPLATCDETSI"

BASE COUNT 775 a 1135 c 1141 g 957 t

ORIGIN

Query Match 100.0%; Score 2331; DB 10; Length 4008;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTTGCAGGGCGAGGGCCCGCGGGCCCGCGCCCGCTCCGGTCTTTCCCTTC 60
DB 76 ATGCTTGCAGGGCGAGGGCCCGCGGGCCCGCGCCCGCTCCGGTCTTTCCCTTC 135

QY 61 CGCGCGCCGCTGCTGCTGCTGCTGCTGCGGATACAAAGCGCCCGGTGCGCGCGC 120
DB 136 CGCGCGCCGCTGCTGCTGCTGCTGCGGATACAAAGCGCCCGGTGCGCGCGC 195

QY 121 GTCCCCCGCTAGTGCCAGAACCTCGCTGCCCATCTCCGAGGCTGACTCTATCTCAC 180
DB 196 GTCCCCCGCTAGTGCCAGAACCTCGCTGCCCATCTCCGAGGCTGACTCTATCTCAC 255

QY 181 CGGTTTGAGCTCTATACGTACAAATTAATCTGCTCTCTCTGTTGATGCTGCCCTCCAC 240
DB 256 CGGTTTGAGCTCTATACGTACAAATTAATCTGCTCTCTCTGTTGATGCTGCCCTCCAC 315

QY 241 ACACATTAGTCGTCGAGGGATAGCATCTTCGCTTTAACCTCCCTCTCTCGGGAA 300
DB 316 ACACATTAGTCGTCGAGGGATAGCATCTTCGCTTTAACCTCCCTCTCTCGGGAA 375

QY 301 AGACCCCGAAGGATCGACTGGATGCTGAGACTCACAGACAGAACTGCAGGAAGAAA 360
DB 376 AGACCCCGAAGGATCGACTGGATGCTGAGACTCACAGACAGAACTGCAGGAAGAAA 435

QY 361 GCAAGAAAGAGGAGCAATGTACAAATTTATCCAGATTCTCGCCATTGTCAATGCCCTCT 420

DB 436 GCAAGAAAGAGGAGCAATGTACAAATTTATCCAGATTCTGCCATTGTCAATGCCCTCT 495
QY 421 CACCTTCCACGTGCGGCACCTTCCTTTGATCGAAGTGGGGTATTTCATGTGTCC 480
DB 496 CACCTTCCACGTGCGGCACCTTCCTTTGATCGAAGTGGGGTATTTCATGTGTCC 555
QY 481 AGTTTCCAGCAGGTTGAAGACTTGAAGCGCCGCGGGAATGTCTTTTTCAGCAGCT 540
DB 556 AGTTTCCAGCAGGTTGAAGACTTGAAGCGCCGCGGGAATGTCTTTTTCAGCAGCT 615
QY 541 CAAACGGTCAGCAGCTGTAAATGGCTGGGGCTCTCTACACGCCCACTGTGAAGACTTC 600
DB 616 CAAACGGTCAGCAGCTGTAAATGGCTGGGGCTCTCTACACGCCCACTGTGAAGACTTC 675
QY 601 CTGGGGACTGAGCCCATCATCTCCGAGCTGGGTGAGCTGAGACTTGAATTCGAACA 660
DB 676 CTGGGGACTGAGCCCATCATCTCCGAGCTGGGTGAGCTGAGACTTGAATTCGAACA 735
QY 661 GAGACCTTGTATCTCTGCTTAATGCTCCAGCTTTTGTGCGAGCTATGCTCTGAGCCCA 720
DB 736 GAGACCTTGTATCTCTGCTTAATGCTCCAGCTTTTGTGCGAGCTATGCTCTGAGCCCA 795
QY 721 GCTGAGTGGGGGATGAAGATGGAGACGATGAATCTTTTTCCTTCAGGAGACTTCC 780
DB 796 GCTGAGTGGGGGATGAAGATGGAGACGATGAATCTTTTTCCTTCAGGAGACTTCC 855
QY 781 CGAGTGTGGACTCTATGAGCGCTCAAGTCCCAAGAGTGGCCGAGTGTGTGCGGG 840
DB 856 CGAGTGTGGACTCTATGAGCGCTCAAGTCCCAAGAGTGGCCGAGTGTGTGCGGG 915
QY 841 GACCTTGGGGGAGGAAGACCTTCAGCAGAGATGGAGACGCTTTCTGAAGCTGACCTG 900
DB 916 GACCTTGGGGGAGGAAGACCTTCAGCAGAGATGGAGACGCTTTCTGAAGCTGACCTG 975
QY 901 CTGTGCCAGGCCCGAGCATGGCCGCGCTCCGGGTTCTGACGGCTATGCGCAGCTT 960
DB 976 CTGTGCCAGGCCCGAGCATGGCCGCGCTCCGGGTTCTGACGGCTATGCGCAGCTT 1035
QY 961 CGGCTCTAGCTTGGAGCGGAACCCCATCTTTTATGGATCTTTTCTCCAGTGGGAA 1020
DB 1036 CGGCTCTAGCTTGGAGCGGAACCCCATCTTTTATGGATCTTTTCTCCAGTGGGAA 1095
QY 1021 GGAGTGCATCTCTGCTGTGTGCTTCCGACCCCAAGACATCGGCGCATGTCTGAAT 1080
DB 1096 GGAGTGCATCTCTGCTGTGTGCTTCCGACCCCAAGACATCGGCGCATGTCTGAAT 1155
QY 1081 GGTCTCTTAGAGACTTAACATGACTGCAACAGGGGACTGCTGTCTATGACACGAG 1140
DB 1156 GGTCTCTTAGAGACTTAACATGACTGCAACAGGGGACTGCTGTCTATGACACGAG 1215
QY 1141 GTGCCCCAGCCCGAGACTGGAGAGTGCATCGCAACAACATGAAGCTCCAGCAGTTTGA 1200
DB 1216 GTGCCCCAGCCCGAGACTGGAGAGTGCATCGCAACAACATGAAGCTCCAGCAGTTTGA 1275
QY 1201 TCTCACTCTCCCTGCCAGACCGCTGCTCACTTTATCAGAGACACCTCTCATGGAC 1260
DB 1276 TCTCACTCTCCCTGCCAGACCGCTGCTCACTTTATCAGAGACACCTCTCATGGAC 1335
QY 1261 AGGCCCCGTGTTCCCGGCTGACGGCCCGCCCTGCTGCTACTACAGATACACCTTATCTC 1320
DB 1336 AGGCCCCGTGTTCCCGGCTGACGGCCCGCCCTGCTGCTACTACAGATACACCTTATCTC 1395
QY 1321 AGAGTCTGTGGCCACAGGCTGACACGCTCTCAGGGAAGAATATGACGTGCTTACCTG 1380
DB 1396 AGAGTCTGTGGCCACAGGCTGACACGCTCTCAGGGAAGAATATGACGTGCTTACCTG 1455
QY 1381 GGGACAGAGGATGGACACTCCACCGGGCTGTGCGCATTTGGAGCTCAGCTCAGTGTCTTG 1440
DB 1456 GGGACAGAGGATGGACACTCCACCGGGCTGTGCGCATTTGGAGCTCAGCTCAGTGTCTTG 1515
QY 1441 CAGGATCTGGCTTGTTCAGACACGCGGTTGAGAGCATGAATTTGACACGAT 1500

1516	GAGGATCTGGCGCTTGTCCCAAGAACACAGCCGGTGTGAGAGCATGAATTTGATCCACGAT	1575
Db		
1501	TGGCTCCTGGTGGGCTCCCATACTGAGTGCACAAAGTGAACACAGCAACTGTGGCGCT	1560
Qy		
1576	TGGCTCCTGGTGGGCTCCCATACTGAGTGCACAAAGTGAACACAGCAACTGTGGCGCT	1635
Db		
1561	CTCCAGAGTGTCTGGAGTGTATCCTTGCCACAGACCCCGGTGTCGGCTTGAGGCTCCGG	1620
Qy		
1636	CTCCAGAGTGTCTGGAGTGTATCCTTGCCACAGACCCCGGTGTCGGCTTGAGGCTCCGG	1695
Db		
1621	CTTGATGCTGTGTGGCCACCGCGGCGAGCACCGCGGGATGTTCAAGATATAGAGTCA	1680
Qy		
1696	CTTGATGCTGTGTGGCCACCGCGGCGAGCACCGCGGGATGTTCAAGATATAGAGTCA	1755
Db		
1681	CGCGATGCTCTTCTTTGTGTCAAAGAACTGGAGAACTCCGCTAGTGTGTTGAAGTT	1740
Qy		
1756	CGCGATGCTCTTCTTTGTGTCAAAGAACTGGAGAACTCCGCTAGTGTGTTGAAGTT	1815
Db		
1741	CGCGTGGTACTGTGGGCGACGTGCTGCTGCTATGTTCCCGACGTTCTGCTGGCGATCC	1800
Qy		
1816	CGCGTGGTACTGTGGGCGACGTGCTGCTGCTATGTTCCCGACGTTCTGCTGGCGATCC	1875
Db		
1801	TGTGTGTGGCACACCCAGTGGAGTGAAGTCTGCTGCTGCTATGTTCCCGACGTTCTGCTGGCGATCC	1860
Qy		
1876	TGTGTGTGGCACACCCAGTGGAGTGAAGTCTGCTGCTGCTATGTTCCCGACGTTCTGCTGGCGATCC	1935
Db		
1861	GTGTGTGTGACCCAGGCGCATGGGGCTTATGCTTCCGAGTCTCAGGAGGTTGGAGCC	1920
Qy		
1936	GTGTGTGTGACCCAGGCGCATGGGGCTTATGCTTCCGAGTCTCAGGAGGTTGGAGCC	1995
Db		
1921	GCCCGCTGGTGGCTTATAGCTTGTGTGGGCGACGACGCGGGACCCCTCAACCCG	1980
Qy		
1996	GCCCGCTGGTGGCTTATAGCTTGTGTGGGCGACGACGCGGGACCCCTCAACCCG	2055
Db		
1981	GCCACACGGTGTGGGGCTGGATGTTGGCTTCTCCTGGGTGTTCTTTCAGCATCC	2040
Qy		
2056	GCCACACGGTGTGGGGCTGGATGTTGGCTTCTCCTGGGTGTTCTTTCAGCATCC	2115
Db		
2041	CTCACTCTCCTGATGTTGGTGCCTCAGCAGCGTGGCGACAGAGGAGCTTCTAGCT	2100
Qy		
2116	CTCACTCTCCTGATGTTGGTGCCTCAGCAGCGTGGCGACAGAGGAGCTTCTAGCT	2175
Db		
2101	AGAGACAAGTGGGCTTAGATCTGGGGCTCCACCTTCTGGGACCAAGCATATAGTCAG	2160
Qy		
2176	AGAGACAAGTGGGCTTAGATCTGGGGCTCCACCTTCTGGGACCAAGCATATAGTCAG	2235
Db		
2161	GACCTCTCCTCTTCGCTGAAGTGAAGCGTGGCCCTGGCGCTGGGTGAAGCGGGC	2220
Qy		
2236	GACCTCTCCTCTTCGCTGAAGTGAAGCGTGGCCCTGGCGCTGGGTGAAGCGGGC	2295
Db		
2221	AGTGGTTTGGTGGCTTCCCTCCACCTTCTGCTGGATCTTCCCAAGCCAGCCAC	2280
Qy		
2296	AGTGGTTTGGTGGCTTCCCTCCACCTTCTGCTGGATCTTCCCAAGCCAGCCAC	2355
Db		
2281	ATCCGGCTCACTGGGCGCCTCTAGCCACGTTGATGAGACCTCCATCTAA	2331
Qy		
2356	ATCCGGCTCACTGGGCGCCTCTAGCCACGTTGATGAGACCTCCATCTAA	2406
Db		
RESULT 2		
AB021291		
LOCUS	4002 bp	linear
DEFINITION	Mus musculus mRNA for semaphorin W, complete cds.	AB021291
ACCESSION	AB021291	ROD 09-APR-1999
VERSION	AB021291.1	GI:4519447
KEYWORDS	semaphorin W,	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
	1 (sites)	
	Encinas, J.A., Kikuchi, K., Chedotal, A., de Castro, F., Goodman, C.S.	
	and Kimura, T.	

QY 358 AAAGGCAAGAGAGAGACGAATGTACAAATTTATCCAGATTCTCGCCATTTCTCAATGCC 417
DB 439 AAAGGCAAGAGAGAGACGAATGTACAAATTTATCCAGATTCTCGCCATTTCTCAATGCC 498
QY 418 TCTCACCCTCTCAGCTGGGACCTTCGCTTTTGATCCGAAGTGGGGGTTATTGATGTG 477
DB 499 TCTCACCCTCTCAGCTGGGACCTTCGCTTTTGATCCGAAGTGGGGGTTATTGATGTG 558
QY 478 TCCAGTTTCCAGCAGGTTGAAAGACTTGAGAGCGCGGGGGAATGTCTTTTGAGCCA 537
DB 559 TCCAGTTTCCAGCAGGTTGAAAGACTTGAGAGTGGCCGGGGAATGTCTTTTGAGCCA 618
QY 538 GCTCAACGCTCAGCAGCTGTAATGCTGGGGCGCTCTCTACACGCCCACTGTGAAGAAC 597
DB 619 GCTCAACGCTCAGCAGCTGTAATGCTGGGGCGCTCTCTACACGCCCACTGTGAAGAAC 678
QY 598 TTCTCTGGGACCTGAGCCCATCATCTCCGAGCTGTGGGTCGAGCTGAGACCTGGATTGGA 657
DB 679 TTCTCTGGGACAGAGCCGATTAATCTCCGAGCTGTGGGTCGAGCTGAGACCTGGATTGGA 738
QY 658 ACAGAGACCTTGTCTATCTCTGCTTAATGCTCCAGCTTTGTGCGAGCTATGTCCTGAGC 717
DB 739 ACAGAGACCTTGTCTATCTCTGCTTAATGCTCCAGCTTTGTGCGAGCTATGTCCTGAGC 798
QY 718 CCAGCTGAGTGGGGGATGAAGATGGAGACGATGAAATCTTTTCTTTCACGGAGACC 777
DB 799 CCGGCTGAGTGGGGGATGAAGATGGAGACGATGAAATCTTTCTTTCACGGAGACC 858
QY 778 TCCGAGCTTGGACTCTATCAGCGCATCAAGTCCCAAGAGTGGCCCGAGTGTGCGG 837
DB 859 TCCGAGCTTGGACTCTATCAGCGCATCAAGTCCCAAGTGGCCCGAGTGTGTCGG 918
QY 838 GGGGACCTTGGGGCAGGAAGACCTTTCAGCAGAGATGACGACGTTTCTGAAGGCTGAC 897
DB 919 GGGGACCTTGGGGCAGGAAGACCTTTCAGCAGAGATGACGACGTTTCTGAAGGCTGAC 978
QY 898 CTGCTGTGCCAGGGGCCAGCATGGCCGGGCTCCGGGGTCTCGAGCTATGCGAGAG 957
DB 979 CTGCTGTGCCAGGGGCCAGCATGGAAGGGCTCGGGGTTCTCGAGGATATGACAGAG 1038
QY 958 CTTCGGCTCAGGCTGGAGCGGGAACCCCATCTTTATGGGATCTTTTCTCCAGTGG 1017
DB 1039 CTTCGGACCTCAGGCTGGCGGGGACCCCTCTTTATGGGATCTTTTCTCCAGTGG 1098
QY 1018 GAAGAGCTGCTCTCTCTGTGTGCTTTCGACCCCAAGACATCCGGGCAAGTGTG 1077
DB 1099 GAAGAGCGCCATTTCTCTGTGTGCTTTCGACCCCAAGACATCCGGGCAAGTGTG 1158
QY 1078 AATGTCTCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCTGTCTATGGAACA 1137
DB 1159 AATGTCTCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCTGTCTATGGAACA 1218
QY 1138 GAGGTGCCCGCCAGACCTGGAGAGTGCATGCGCAACATGAAGCTCCAGCATTT 1197
DB 1219 GAGGTGCCCGCCAGACCTGGAGAGTGCATGCAACAAATGAAGTTCAGCATTT 1278
QY 1198 GGATCTCACTCTCTCTGCGACACCGCTGCTCACCCTTTATCAGAGCACCCCTCTCATG 1257
DB 1279 GGATCTCACTCTCTCTGCGACACCGCTGCTCACCCTTTATCAGAGCACCCCGCTCATG 1338
QY 1258 GACAGCCCGTGTTCGGGCTGACGGCCGCCCGCTGCTGCTCACTACAGATACAGCCAT 1317
DB 1339 GACAGCCCGTGTTCAGCTGATGGCCGCCCGCTGCTGCTCACTACGATACAGCCAT 1398
QY 1318 CTCAGAGTCTGGCCACAGGCTGACCGCTCTCAGGAAAGATATACGCTGCTCTAC 1377
DB 1399 CTCAGAGTGTGGCTCACCGGGTGAACGCTCTCAGGAAAGATATATGCTGCTCTAC 1458
QY 1378 CTGGGACAGAGGATGGACACCTTCCACCGGGTGTGGCATTTGGAGCTCAGCTCAGTGTG 1437
DB 1459 CTAGGACAGAGGATGGGACCTTCCACCGGGTGTGGCATTTGGAGCTCAGCTCAGTGTG 1518
QY 1438 TTGGAGGATCTGGCCCTGTGTCCAGAACACAGCCGGTTTGAGAGCATGAAATGTGACCAC 1497

DB 1519 CTGGAGGATCTGGCCCTGTTCCTCGGAACCAACCGTTTGAGAGCATGAAATGTACCAT 1578
QY 1498 GATTGGCTCTCTGGTGGCTCCCATCTGAGGTGACAAAGTGAACACCAAGCAACTGTGGC 1557
DB 1579 GATTGGCTCTCTGGTGGCTCCCATCTGAGGTGACAAAGTGAACACCAAGCAACTGTGGC 1638
QY 1558 CGTCTCCAGAGCTGCTCGGAGTGTATCTCTGGCCAGGACCCGCTGTGCGCCTGGAGCTTC 1617
DB 1639 CGTCTCCAGAGCTGCTCAGAATGTATCTCTGGCCAGGATCCGCTGTGCTCTGGAGCTTC 1698
QY 1618 CGGCTGATGCTGTGTGGCCACCGCGGAGCAGCCGGGGATGGTTCAAGATATAGAG 1677
DB 1699 CGGCTCGATGCTGTGTGGCCATCGAGCGCAGCACCGTGGGATGGTTCAAGACATAGAG 1758
QY 1678 TCAGCGGATGCTCTCTCTTTGTCCTCAAAAGAACCTGGAGAACATCCCTAGTGTGTTGAA 1737
DB 1759 TCAGCGATGCTCTCTCTTTGTCCTCAAAAGAACCTGGAGAACATCCGCTAGTGTGTTGAA 1818
QY 1738 GTTCGGGTGGCTACTGTGGGCCACGTTGCTCTGCCATGTTCCGCCAGTTCTGCCTGGCA 1797
DB 1819 GTTCCTGTGGCTACGTTGGGCCACGTTGCTCTGCCGTGCTCCGCCAGTTCTGCCTGGCA 1878
QY 1798 TCTGTGTGTGCACAGCCAGTGCAGTGCCTGCTCCTCCCGAGGATGGACTA 1857
DB 1879 TCTGTGTGTGCACAGCCAGTGCAGTGCCTGCTCCTCACCCTCCGAAAGGATGGACTA 1938
QY 1858 GAGGTGGTGTGACCCAGGGGCCATGGGGCTTATGCTTGGAGTGTTCAGGAGGTGGA 1917
DB 1939 GAGGTGGTGTGACCCAGGGGCCATGGGGCTTATGCTTGGAGTGTTCAGGAGGTGGA 1998
QY 1918 GCCGCCGGTGGTGGCTCTTATAGCTTGTGTGGGCGAGCAGCGGGACCCCTCAAAAC 1977
DB 1999 GCCGCCGGTGTGGCTCTTATAGCTTGTGTGGGCGAGCAGAGGGGGCCCTCAAAAC 2058
QY 1978 CGGGCCACACGTTGTGGGGCTGAGTGGTGGCTTCTCCTGGGTGTTCTTGACGCA 2037
DB 2059 CGGGCCACACGTTGTGGGGCTGAGTGGTGGCTTCTCCTGGGTGTTCTTGACGCA 2118
QY 2038 TCCCTCACCTCTCTCTGATTGTCGCCCTCAGCAGCTCGGCGACAGAGGGAGCTTCTA 2097
DB 2119 TCTCTACCTCTCTCTGATTGTCGCCCTCAGCAGCTCGGCGAGCAGAGGGAGCTTCTA 2178
QY 2098 GCTAGAGACAAGTGGGCTTAGATCTGGGGCTTCCACCTTCTGGGACCAAGCTATAGT 2157
DB 2179 GCTAGAGACAAGTGGGCTTAGATCTGGGGCTTCCACCTTCTGGGACCAAGCTATAG 2238
QY 2158 CAGGACCTCTCTCTCTGCTGAGATGACGGCTGCCCTGGCCCTGGTAAAGCG 2217
DB 2239 CAAGACCTCTCTCTCTGCTGAGATGAAGCTGAAGCTGCCCTGGCCCTGGTAAAGCG 2298
QY 2218 GGCAGTGGTTTGGTGGCTTCCCTCCACCTTCTCTGCTGATTCTTGGCCCAAGCCAGCC 2277
DB 2299 GGCAGTGGTTTGGTGGCTTCCCTCCACCTTCTCTGCTGATTCTTGGCCGAGCCAGCC 2358
QY 2278 CACATCCGGCTCAGTGGGGCTCTAGCCACCTGTGATGAGACCTCCATCTAA 2331
DB 2359 CACATCCGGCTCAGTGGGGCTCTCTCTGGCCAGCTGCGATGAGACGCTATCTAA 2412

RESULT 3

AF038652
LOCUS 2354 bp mRNA linear ROD 24-MAR-2003
DEFINITION Mus musculus semaphorin M mRNA, partial cds.
ACCESSION AF038652
VERSION AF038652.1 GI:4104674
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2354)
AUTHORS Jang,W., Spilison,S.V., Hua,A., Roe,B. and Meisler,M.H.

TITLE Large-scale comparative sequence analysis of human and mouse
JOURNAL genomic DNA reveals coding regions of three new genes
REFERENCE 2 (bases 1 to 2354)
AUTHORS Jang, W., Spillson, S.V. and Meisler, M.H.
TITLE Direct Submission
JOURNAL Submitted (12-Dec-1997) Human Genetics, University of Michigan,
4708 MS II, Box 0618, Ann Arbor, MI 48109-0618, USA
FEATURES
Location/Qualifiers
1..2354
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/chromosome="6"
/map="between D6Mit164 and D6Mit128"
<1..2171
/note="Semam; member of semaphorin gene family"
/codon_start=3
/product="semaphorin M"
/protein_id="AAD02103.1"
/db_xref="GI:4104675"
/translation="ADSYLIRFAAPHTYNSALLYVDPAHFLYVYVARDISFALILPES
GERPRIDMWPETHQRNCRKKRKEDECHNFQILAIANASHLLTCTGFADPKGV
IDVSFQOERLESGRGKCPPEPAQSAANMAGVLYLATYKFLGTEPIISRAVGR
EDWIRTEGLSSWLNAPFAVAAMVLSPAEWGDEGDEIFFEFTETSRVLDSEYRIKVP
RVARVCADLGGRTLAQRTWTLFLKADLLCPGPEHGRASGVLDQWTELRPOPGAGTPL
FYGIFSSOWEAGATLSAVCAPRPQDIRAVLNGPFRKELHDCNRGLPVMNDNEYPORPGE
CITNNMFQOQSGSLSLPDRVLTFRIDRPLMDRPFVADGRPLLVTDYALRVVHR
VTSUGKEIDVLYLGTEDEGHLRAVRIGLQSLVEDIALFPETQFVPMKLYHDLVLY
GSHTEPQVNTSGNRVSCSECLADPVCAMSFRLDACVAHAGEHRTGMVDIESAD
VSSLCPEKPGHVVVFVPAVTVGHVLPVPCSPSSAWSCVWHQPSGVTSLPFRDGL
VVVTPGAMGAVACQCEQGAARVVAAYSLVNGSQRPANRAHTVVGAGLVGFFLGLA
ASLPLLLIGRQRRORRELLARDKVLGLCAPSGTTSYSQDPPSPSPEDERLPLAL
GKRSGGGFFPPPELLDSCPSPAHRLTGAPLAYCDETSI"
BASE COUNT 462 a 661 c 652 g 539 t
ORIGIN
Query Match 86.0%; Score 2004..6; DB 10; Length 2354;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 2067; Conservative 0; Mismatches 104; Indels 0; Gaps 0;
QY 161 AGGTGACTCTATCTACCCGGTTTGAGCGCTCTACATACGATACAAATTTACTCTGCTTCC 220
DB 1 AGGTGACTCTATCTACCCGGTTTGAGCGCTCTACATACGATACAAATTTACTCTGCTTCC 60
QY 221 TTGTGGATCTTCCACACACTTTACGTCGGTGACGGGATAGCATCTCGCTTAA 280
DB 61 TGGTGGATCTTCTCACACACTTTACGTCGGTGACGGGATAGCATCTCGCTTAA 120
QY 281 CCCTCCCTTCTCTGGGAAAGACCCCGAAGGATCGACTGGATGCTACCTGAGACTACA 340
DB 121 CCCTGCCCTTCTCTGGGAAAGCTCTCAAGATGCTGATGATGTCGCCGAGACTACA 180
QY 341 GACAGAACTGCAGGAAGAAAGCAAGAAAGAGGAGCAATGTACAAATTTATCCAGATTC 400
DB 181 GACAGAACTGTAGGAAGAAAGCAAGAAAGAGGAGCAATGTACAAATTTATCCAGATTC 240
QY 401 TCGCCATGTCAATGCCCTCTCACCTCTCAGTCGGCGCACCTTCGCTTTTGTATCCGAAGT 460
DB 241 TCGCCATGTCAATGCCCTCTCACCTCTCAGTCGGCGCACCTTCGCTTTTGTATCCGAAGT 300
QY 461 CGGGGGTTATTGATGTGTCAGTTTCCAGAGGTTGAAGACTTCAGAGCGCGCGGGGA 520
DB 301 CGGGGGTTATTGATGTGTCAGTTTCCAGAGGTTGAAGACTTCAGAGGTTGCCCGGGGA 360
QY 521 AATGTCTTTTTCAGCAGCTCAACGGTCAGCAGCTGTAAATGGCTGGGGCGCTCTCTTACA 580
DB 361 AATGTCTTTTTCAGCAGCTCAACGGTCAGCAGCTGTAAATGGCTGGGGCGCTCTCTTACA 420
QY 581 CGGCCACTGTGAAGAACTCTTGGGGACTGAGCCCATCATCTCCCGAGCTGTGGGTCGAG 640
DB 421 CGGCCACTGTGAAGAACTCTTGGGGACTGAGCCCATCATCTCCCGAGCTGTGGGTCGAG 480

QY 641 CTGAGGACTGGATTGGAACAGAGAGACCTTGTCTCATCTCTGGCTTAATGCTCCAGCCTTTGTCCG 700
DB 481 CTGAGGACTGGATTGGAACAGAGAGACCTTGTCTCATCTCTGGCTTAATGCTCCAGCCTTTGTCCG 540
QY 701 CAGCTATGGTCTTGAGCCAGCTGAGTGGGGGATGAAGATGGAGACGATGAATCTTTT 760
DB 541 CAGCTATGGTCTTGAGCCCGGCTGAGTGGGGGATGAAGATGGAGACGATGAATCTTTCT 600
QY 761 TTTTCTTCCAGGAGACCTCCGAGTGTGGACTCTCTATGAGGCAATCAAGTCCCAAGAG 820
DB 601 TTTTCTTCCAGGAGACCTCCCGAGTCTTGGACTCATATGAGGCAATCAAGTCCCAAGAG 660
QY 821 TGGCCGAGTGTGTGGGGGACCTTTGGGGGAGGAGAACCTTTCAGCAGAGATGAGCA 880
DB 661 TGGCCGAGTGTGTGGGGGACCTTTGGGGGAGGAGAACCTTTCAGCAGAGATGAGCA 720
QY 881 CGTTTCTGAAGGCTGACCTGTGTGCCAGGGCCGAGCATGAGCGGGCTCCGGGGTTC 940
DB 721 CGTTTCTAAGGCTGACCTGTGTGCCAGGGCCGAGCATGAGCGGGCTCCGGGGTTC 780
QY 941 TGCAGGCTATGCGAGAGCTTCGGCTCAGCCTGAGCGGGAACCCCATCTTTTATGGGA 1000
DB 781 TGCAGGATATGACAGAGCTTCGACCTCAGCCTGGCGGGGACCCCTCTTTTATGGCA 840
QY 1001 TCTTTTCTCCAGTGGGAAGAGCTGCCATCTCTGCTGTGTGCTGTGCTGTGCTTCCGACCCCAAG 1060
DB 841 TCTTTTCTCCAGTGGGAAGAGCGCCCATTTCTGCTGTGCTGTGCTTCCGACCCCAAG 900
QY 1061 ACATCCGGGAGTGTGTAATGGTCTTTAGAGACTAAACATGACTGCAACAGGGAC 1120
DB 901 ACATCCGGGAGTGTGTAATGGTCTTTAGAGACTAAACATGACTGCAACAGGGAC 960
QY 1121 TGCTGTCTATGAGAACAGGCTGCCCCAGCCAGACCTGGAGAGTGCATCGCAACAACA 1180
DB 961 TACCTGTATGGAACAACAGGCTGCCCGCCAGACCTGGAGAGTGCATCACAACAACA 1020
QY 1181 TGAAGCTCCAGAGTTTGGATCCTCCTCCTGCCAGACCGCTGTGCTACCTTTATCA 1240
DB 1021 TGAAGTTCCAGAGTTTGGATCCTCCTCCTGCCAGACCGCTGTGCTACCTTTATCA 1080
QY 1241 GAGACACCTCTCATGAGACAGGCCGCTGTTCGGGGTGACGGCGGCCCTCTGCTGCTCA 1300
DB 1081 GAGACACCTCTCATGAGACAGGCCGCTGTTCGGGGTGACGGCGGCCCTCTGCTGCTCA 1140
QY 1301 CTACAGATACAGCCTATCTCAGAGTGTGGCCACAGGCTGACAGCCTCTCAGGGAAG 1360
DB 1141 CTACGGATACAGCCTATCTCAGAGTGTGGCTCACCGGTGACAGCCTCTCAGGGAAG 1200
QY 1361 AATATGAGTGTCTTACCTGGGGACAGAGGATGACACCTCCACCGGGCTGTGCGCATTC 1420
DB 1201 AATATGAGTGTCTTACCTGGGGACAGAGGATGACACCTCCATCGGGCTGTGCGCATTC 1260
QY 1421 GAGCTCAGCTCAGTGTCTTGGAGATCTGGCCCTTGTTCAGAAACACAGCCGGTGTAGA 1480
DB 1261 GAGCTCAGCTCAGTGTCTTGGAGATCTGGCCCTTGTTCGGGAAACACAGCCGGTGTAGA 1320
QY 1481 GCATGAATTTTACCAACGATTTGGCTCTCTGGGGCTCCCATCTCAGGTGACACAGTCA 1540
DB 1321 GCATGAATTTTACCAACGATTTGGCTCTCTGGGGCTCCCATCTCAGGTGACACAGTCA 1380
QY 1541 ACACAGCAACTGTGGCGCTCTCCAGAGCTGTCTGGAGTGTATCTCTGGCCAGGACCCCG 1600
DB 1381 ACACAGCAACTGTGGCGCTCTCCAGAGCTGTCTCAGATGTATCTCTGGCCAGGATCCCG 1440
QY 1601 TGTGGCCTGGAGCTTCCGGCTTGTGCTGTGGCCACCGCCGCGAGCACCCTGGGA 1660
DB 1441 TGTGTGCTGGAGCTTCCGGCTGATGTGTGTGGCCCATCGACGGAGCACCCTGGGA 1500
QY 1661 TGGTTCAAGATATAGAGTCAGGGATGTCTCTTTTGTGTCCAAAGAACCTTGAGAAC 1720
DB 1501 TGGTTCAAGATATAGAGTCAGAGATGTCTCTTTTGTGTCCAAAGAACCTTGAGAAC 1560

Db 336 AGACCCCGCAGATTGACTGGATGGTTCTCAGGCTCACAGACAACTGTAGGAAGAAA 395
 QY 361 GCGAAGAAGAGGACGAATGTACAAATTTTATCAGATTTCTCGCATTTGTCAATGCCCTCT 420
 Db 396 GCGAAGAAGAGGACGAATGTACAAATTTTGTCCAGATTTCTCGCATTTGCCAATGCCCTCT 455
 QY 421 CACCTCTCAGCTGGGCGACCTTCGCTTTTATCGGAAGTGGGGTTATTTGATGTGTC 480
 Db 456 CACCTCTCAGCTGGGCGACCTTCGCTTTTATCGGAAGTGGGGTTATTTGATGTGTC 515
 QY 481 AGTTTCCAGCAGTTTGAAGACTTGAGAGCGCGGGGAAATGTCCTTTTGAAGCAGCT 540
 Db 516 AGTTTCCAGCAGTTTGAAGACTTGAGAGTGGCGGGGAAATGTCCTTTTGAAGCAGCT 575
 QY 541 CAACGGTCAGCAGCTGAATGGCTGGGGGCTCTCTACACGGCAGCTGTGAAGAACTTC 600
 Db 576 CAGCGGTGAGCAGCTGAATGGCTGGGGGCTCTCTATGTCGACACTGTCAAAACTAC 635
 QY 601 CTGGGGACTGAGCCATCATCTCCGAGCTGTGGGTGAGCTGAGGACTGATTCGAACA 660
 Db 636 CTGGGGAGCGGACCAATATCACCAGACAGCTGGGTGCTGCGGAGACTGATTCGGACA 695
 QY 661 GAGACCTTGTCTGCTTAATGCTCCAGCTTTGTCGAGCTATGCTCCTGAGCCCA 720
 Db 696 GATACCTTGCCTTCTGCTGACGCCGCCGCTTTGTCGAGCGCTGGCTTTGAGCCCA 755
 QY 721 GCTGAGTGGGGGATGAAGATGGAGACGATGAAATCTTTTCTTTCACGGAGACCTCC 780
 Db 756 GCCGAATGGGGGATGAAGATGGAGACGACGAAATCTACTTCTTTACGGAGACTCC 815
 QY 781 CGAGTGTGGACTCTATGACCGCATCAGTCCAGATGAGCTGGCCGAGTGTGCGGG 840
 Db 816 CGAGCATTTGACTCATAGAGCGCATTAAGTCCACCGGTGGCCGCTGTGTGCGGG 875
 QY 841 GACCTTGGGGCAGAGACCTTTCAGCAGATGGAGAGCTTTCGAAGCTGACCTG 900
 Db 876 GACCTGGGGCCGAGAGACCTTCAGCAGATGGAGAGCTTTTGAAGCTGACCTG 935
 QY 901 CTGTGCCAGGGCCGAGCATGGCGGGCTCCGGGTTCTGAGGCTATGGCAGAGCTT 960
 Db 936 CTGTGTCCAGGGCTGAGCATGGCGGGCTCCAGTGTCTGCAGGATGTTGCTGTGCTT 995
 QY 961 CGGCTCAGCTGAGCGGGAAACCCCATCTTTTATGGGATCTTTCTCCAGTGGGAA 1020
 Db 996 CGACCTGAGCTGGGGCAGGAGCTCCCATCTTTTATGGCATCTTTCTCCAGTGGGAG 1055
 QY 1021 GGAGCTGCATCTCTGTGTGTGCTTCCGACCCCAAGACATCCGGGAGTGTGTAAT 1080
 Db 1056 GGGCTACTATCTCTGTGTGTGCTTCCGACCAAGACATTCGGACAGTGTGTAAT 1115
 QY 1081 GGTCCCTTTAGAGAGCTAAACATGACTGCAACAGGAGCTGCCTGTGATGGACAACGAG 1140
 Db 1116 GGTCCCTTCAGAGAACTAAACATGACTGCAACAGAGAGCTGCCTGTGTAATGAT 1175
 QY 1141 GTGCCCCAGCCAGCTGGAGATGATCCGCAACACATGAAGCTCCAGCTTTGGA 1200
 Db 1176 GTGCCCCAGCCAGCTGGAGATGATCACCACCAACATGAAGCTCCGAGCTTTGGC 1235
 QY 1201 TCCTCACTCTCCCTGCGAGACCGCTGCTACCTTTTATCAGAGACCCCTCTCATGGAC 1260
 Db 1236 TCATCTCTCTCCCTGACCGGTACTCACTTTCATCCGGGACCCCTACTCATGGAC 1295
 QY 1261 AGGCCCTGTTTCCGGCTGAGCGGCCCTCTGCTGCTACTACAGATACAGCTATGTC 1320
 Db 1296 AGGCCAGTGTTCAGCTGATGGCCACCCCTGCTGCTACTACAGATACAGCTATGTC 1355
 QY 1321 AGAGTCTGGCCACAGGCTGACAGCTCTCAGGGAAGAAATATGAGCTGCTTACTG 1380
 Db 1356 AGAGTCTGGCCACAGGCTGACAGCTCTCAGGGAAGAAATATGAGTGTGCTTACTG 1415
 QY 1381 GGGACAGAGGATGACACCTCCACCGGCTGTGCGCATTTGAGCTCAGCTGCTGTG 1440
 Db 1416 GGGACAGAGGATGACACCTCCACCGAGAGTGGGATCGGAGCTCAGCTCAGCTGCTT 1475

QY 1441 GAGGATCTGGCTTGTTCACAGAACCCAGCGGTTGAGAGCATGAAATTTGACCACGAT 1500
 Db 1476 GAAGATCTGGCTTATTCACAGAGCCACAGCCAGTTGAGAACATGAATTTGACCACAGC 1535
 QY 1501 TGGCTCTGGTGGGCTCCCATACTAGAGTGACAAAGTGAACACAGCAACTGTGGCCGT 1560
 Db 1536 TGGCTCTGGTGGGCTCCCATACTAGAGTGACAAAGTGAATACAACTGTGGCCGT 1595
 QY 1561 CTCAGAGCTGCTCGGAGTGATCTGCGCCAGAGACCCGCTGCGGCTGGAGCTTCGG 1620
 Db 1596 CTCAGAGCTGCTCAGAGTGATCTGCGCCAGAGACCCAGTCTGTGCTGGAGCTTCGG 1655
 QY 1621 CTGATGCTTGTGGGCCACCGCGGAGCACCGCGGATGGTTCAAGATATAGAGTCA 1680
 Db 1656 CTGATGAGTGTGGGCCATCCGCGGAGCACCGAGGTTGGTCCANAGACATAGAGTCA 1715
 QY 1681 CGGATGCTCTTCTTTGTGCCAAAGAACTTGAGAACATCCCGTGTGTTGAAGTT 1740
 Db 1716 GCAGATGCTCTCTTGTCTTAAGAGCTGGAGAACGTTCCAGTAGTGTGTTGAAGTT 1775
 QY 1741 CCGTGGCTACTGTGGGCCAGCTGCTGCATGTTCCCGAGTCTGCTGGGATCC 1800
 Db 1776 CCGTGGCTAGCTGCGCATGTGCTTGGCATGTTCTCAAGCTCAGCATGGGATCC 1835
 QY 1801 TGTGTGGCACCCAGCCAGTGGAGTGACTGCGCTCACTCCCGGAGGATGGAGTAG 1860
 Db 1836 TGTGTGGCACCCAGCCAGTGGAGTGACTGCACTCACCCCGGAGGATGGAGTAG 1895
 QY 1861 GTGTGTGACCCAGGCGCATGGGCTTATGCTTCGAGTGTGAGAGGTTGAGGCT 1920
 Db 1896 GTGTGTGACCCAGGCGCATGGGCTTATGCTTGAATGTGAGAGGTTGGGCT 1955
 QY 1921 GCCCGTGGTGGCTTATAGCTTGTGGGCGAGCCAGCGGAGGACCTCAACCGG 1980
 Db 1956 GCCCATGTGTAGAGCTTACAGCTTGTATGGGCGAGCCAGAGATGCTCCGAGCGG 2015
 QY 1981 GCCCACACCGTGTGGGCTGGATGGTGGCTTCTCTCGGTTGTTCTTCGAGCATCC 2040
 Db 2016 GCCCACAC- - -AGTGGGGCGGAGTGGCTGGCTTCTTCTGGGATCTCGCAGCATCC 2072
 QY 2041 CTCACTCTCTCTGATTTGCTGCGCTCAGAGCGTCCGCGACAGAGGAGCTTCTAGCT 2100
 Db 2073 CTGACTCTCATTTGCTGCGCTGCGCTCAGAGCGAGCGGACAGAGGAACTTCTGCT 2132
 QY 2101 AGAGCAAGGTGGCTTAGATCTGGGCTCCACTTCTGCGACACAGCTATAGTCAG 2160
 Db 2133 AGAGCAAGGTGGCTTAGATCTGGGCTCCACTTCTGCGACACAGCTATAGTCAG 2192
 QY 2161 GACCTCTCTCTCTCTGAGATGAAGCTGCGGCTGCGGCTGGGCTGGGTAAGCGGCG 2220
 Db 2193 GACCTCTCTCTCTCTGAGATGAAGCTGCGGCTGCGGCTGGGCTGGGTAAGCGGCG 2252
 QY 2221 AGTGGTTTGTGGCTTCCCTCCAGCTTCTGCTGATTTCTGCGGAGGAGGAGGAG 2280
 Db 2253 AGTGGTTTGTGGATTTCTCAGGCTTCTGCTGATTTCTGCTGATTTCTGCGGAGGAGGAG 2312
 QY 2281 ATCGGCTCACTGGGCGGCTTACGACAGTGTGATGAGACTCCATCTA 2330
 Db 2313 ATTCGGCTAACTGGGCTTCTTACCCACATGTGATGAACATCCATCTA 2362

RESULT 5
 AX136143
 LOCUS AX136143 2971 bp DNA linear PAT 30-MAY-2001
 DEFINITION Sequence 65 from Patent EP1067182.
 ACCESSION AX136143
 VERSION AX136143.1 GI:14272551
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE	1	Ota,T., Isogai,T., Nishikawa,T., Kawai,Y., Sugiyama,T. and Hayashi,K.
AUTHORS		Secretory protein or membrane protein
TITLE		Patent: EP 1067182-A 65 10-JAN-2001;
JOURNAL		Helix Research Institute (JP)
FEATURES		Location/Qualifiers
source		1..2971
		/organism="Homo sapiens"
		/mol_type="genomic DNA"
		/db_xref="taxon:9606"
CDS		103..2415
		/note="unnamed protein product"
		/codon_start=1
		/protein_id="CAC39714.1"
		/db_xref="GI:1427252"
		/translation="MPASARPRPGQPQTASPFLLLLAVLSVPSGRVPRSPRTS LPISADSLTRFAPHYNTYSLVDFASHLYVGARDTIFALSLPSGERPRIDW MYPEAHRONCRKKKEDECHNFVGLIAANASHLLTCTGFADPKCVIDVSRFQY ERLESGRKCPPEPAQSAAMVAGVLYAATVKNYLTGETPIITRAVGAEDWIRDTL PSWLNAPAFVAVALSPAEMWGEDDEIDYFFETTSRAFDSYERIKVPRVARVCAGD LGKRTLOQRWTFLLKADLDCPPEHGRASSVLODVAVLRPELGAGTPIFYGFSSOW EGATISAVCAPRPQDIRVNLNPGPFRELKHDNRLPVVDNDVQPPRGECLTNKKLR HRGSLSLDRVLTFIRBHPMDRPFVPADGHPLLVTDYTLRVAHVRSLSKEY DVLYLGTEGHLHRAVRIPGAQLSVLEDLALPEPQPVVENMKLYHSLWLSRTEVQY NTNCGRLQSCSECLAQDPVCAMSFRLDECVAHAGEHRLVQDIESADVSLCPKEP GERPVYFVPVATAARVLPSPSSAMASCVHQPSGVLTALTPRRDGLSEVYVTPGAMG AYACBCQBGGAHVAAVYSLVNGSORDAPSRHAHTVAGLAGLGFLLGILAAASLTLLIGR ROORRORRELLARDVGLDGLAPPSSGTSYSODPPSPSPEDERLPLALAKRSGSGGF SPFFLLDPCPSAHLRLTGAFLATQDETSI"
BASE COUNT	582 a	853 c 858 g 678 t
ORIGIN		
Query Match	78.4%;	Score 1827.6; DB 6; Length 2971;
Best Local Similarity	87.3%;	Pred. No. 0;
Matches 2033;	Conservative	0; Mismatches 279; Indels 18; Gaps 2;
QY	1	ATGCTTTGCAGGGCGAGCGGCCCGCGGGCCCCCGGCCCTCCGGTCTTTTCCCTTC 60
DB	103	ATGCGGGGCTCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 162
QY	61	CCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
DB	163	CC-----GCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 207
QY	121	GTCCCGCGCTCAGTCCAGAACCTCGCTGCGCATCTCGAGGCTGACTCTCTACCC 180
DB	208	GTCCCGCGCTCGTGCCAGACCTCGCTTCAAATCTGTAGGCTGACTCTGTCTCACC 267
QY	181	CGGTTTGCAGCGTCTACGTACAAATTAATCTGCTGCTCTCTCTGTTGGATCGCTCCAC 240
DB	268	CGGTTTGCAGCTCCCTCACACATACAAATTAATCTGCTGCTCTCTGTTGGATCGCTCCAC 327
QY	241	ACACTTTAGTTCGGTGCAGGGATAGCATCTTCGCTTTAAACCTCCCTCTCTCGGNA 300
DB	328	ACACTTTAGTTCGGCGCGCGGACACCACTTCCTGCTTTATCCCTGCTCTCTCGAGGAG 387
QY	301	AGACCCGGAAGGATCGACTGGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB	388	AGACCCGCGAGGATTCAGTGGATGTTCTCTGAGGCTCACAGACAGAACTGTAGGAAGA 447
QY	361	GGCAAGAAAGAGGAGATGTACAAATTTATTCAGATTCGCGCAATGTCAATGCCCTCT 420
DB	448	GGCAAGAAAGAGGAGGAAATGTACAAATTTGTTCAGATTCGCGCAATGTCCAAATGCCCTCT 507
QY	421	CACCTTCCTCAGCTGCGGACCTTCGCTTTTGTATCCGAAGTTCGGGTTATGTATGTGCC 480
DB	508	CACCTTCCTCAGCTGCGGACCTTCGCTTTTGTATCCGAAGTTCGGGTTATGTATGTGCC 567
QY	481	AGTTTCCAGCAGGTTGAAGACTTCAGAGCGCGCGGGGGAATGTCTTTTTCAGCCAGCT 540
DB	568	AGTTTCCAGCAGGTTGAAGACTTCAGAGTGCCTCGCGGGGGAATGTCTTTTTCAGCCAGCT 627

QY	541	CAACGGTCAGCAGCTGTAAATGGCTGGGGCGCTCTCTACACCGCCACTGTGAAGAACTTC 600
DB	628	CAGCGGTACGACGCTGTAAATGGCTGGGGGGTCTCTATATGCTGCCACTGTGAAAACTAC 687
QY	601	CTGGGGAGTACGCCCATCATCTCCGAGCTGTGGGTGAGCTGAGACCTGGATTTCGAACA 660
DB	688	CTGGGGAGCGGAGCAATTAATCACAGAGCAGTGGGTCTGTCGGAGGACTGGATTTCGGACA 747
QY	661	GAGACCTGCTCATCTGCTGCTTAAATGCTCCAGCTTTTGTCCGAGCTATGCTCTGAGCCCA 720
DB	748	GATACCTTGCCCTTCCCTGCTGAACGCCCGAGCTTTGTCCGAGCCGTGGCTTGAGCCCA 807
QY	721	GCTGAGTGGGGGATGAAGATGGAGACGATGAAATCTTTTTTTTCTTCCAGGAGACCTCC 780
DB	808	GCCGAATGGGGGATGAAGATGGAGACGAGAACTACTTCTTCTTTACGGAGACTTCC 867
QY	781	CSAGTGTGGACTCTCTATGAGGCGATCAAGTTCACAGTCCCAAGACTGCCCGAGTGTGGGG 840
DB	868	CGAGCATTTGACTCATACGAGCGCATTAAGTCCACGGGTGGCGGTGTGTGGGG 927
QY	841	GACCTTGGGGGAGGAAGACCTTCAGCAGAGATGGACGACGTTTCTGAAGCTGACCTG 900
DB	928	GACCTCGGGGGCGGAAGACCTTCAGCAGAGATGGACGACGTTTGTGAAGCTGACCTG 987
QY	901	CTGTGCCAGGCCCGAGCATGCGCGGCTCCGGGTTCTCGAGCTATGCGAGACTT 960
DB	988	CTCTGTCCAGGGCTTGAGCATGGCGGGCTCCAGTGTCTCGAGGATTTGTGTGCTT 1047
QY	961	CGGCTCAGCTGGAGCGGGAACCCCACTTTTATGGATCTTTTCCCTCCAGTGGGA 1020
DB	1048	CGACCTGAGCTTGGGGCAGGGACTCCCATCTTTTATGGCATCTTTTCTTCCAGTGGAG 1107
QY	1021	GGAGCTGCTCTCTGCTGCTTCCGACCCCAAGACATCCGGGACGTGCTGAAT 1080
DB	1108	GGGGTACTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1167
QY	1081	GGTCCCTTTAGAGACTAAACATGACTGCAACAGGGGACTGCCCTGTATGACAAACAG 1140
DB	1168	GGTCCCTTTAGAGACTAAACATGACTGCAACAGGGGACTGCCCTGTGTCGAGCAATGAT 1227
QY	1141	GTGCCCCAGCCAGACCTGGAGAGTGCATCGCAACACATGAAGCTCCAGCAATTTGA 1200
DB	1228	GTGCCCCAGCCAGACCTGGAGAGTGCATCACAACATGAAGCTCCGGACATTTGGC 1287
QY	1201	TCTCCTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
DB	1288	TCATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1347
QY	1261	AGGCGCGTGTCCCGGTGACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
DB	1348	AGGCGCGTGTTCAGCTGATGGCCACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1407
QY	1321	AGAGTGTGGGCCACAGGCTGACCGCTCTCAGGAAAGATATGACGCTCTACCTG 1380
DB	1408	AGAGTGTGGGCCACAGGCTGACCGCTCTCAGGAAAGATATGATGCTGCTTACCTG 1467
QY	1381	GGGACAGAGGATGGACCTCCACCGGCTGTGCGCATTTGGAGCTCAGCTGAGTCTTGTG 1440
DB	1468	GGGACAGAGGATGGACCTCCACCGGAGTGGGATCGGAGCTCAGCTCAGCGTCTT 1527
QY	1441	GAGGATCTGGCTTGTTCAGAACACAGCGGTTGAGAGCATGAAATTTGACACGAT 1500
DB	1528	GAAGATCTGGCTTATTCAGAGCCACAGCGGTTGAGAACATGAAATTTGACACAG 1587
QY	1501	TGGCTCTGGTGGCTCCCATACTGAGGTGACAAAGTGAACACAGCAACTGTGGCGCT 1560
DB	1588	TGGCTCTGGTGGCTCCCGTACTGAGGTGACAAAGTGAATCAACCAACTGTGGCGCT 1647
QY	1561	CTCCAGAGCTGCTGGAGTGTATCTTGGCCCAAGGACCCCGCTTGGCGCTTGGAGTCCGG 1620
DB	1648	CTCCAGAGCTGCTCAGAGTGTATCTTGGCCCAAGGACCCAGTCTGTGCTTGGAGTCCGG 1707
QY	1621	CTTGATGCTTGTGTGGCCACCGCGGAGACCCCGGCGATGGTTCAAGATATAGATCA 1680

Db 1708 CTGGATGAGTGTGGGCCATGCCGGGAGCACCGAGGGTTGGTCCAAGACATAGAGTCA 1767
Qy 1681 GCGGATCTCTCTTTCTGCCAAAGAACCTGGAGAACATCCCTAGTGTGTGAAGTT 1740
Db 1768 GCAGATGTCTCTTTGTGTCTTAAAGAGCTGGAGAACCTCCAGTAGTGTGTGAAGTT 1827
Qy 1741 CCGTGGCTACTGTGGGCCACGTGTCTGCCATGTTCGCCAGTTCTGGCTGGGCATCC 1800
Db 1828 CCCGTGCTACAGCTGCGCATGTGTCTTGCCATGTCTTCCAAGCTCAGCATGGGCATCC 1887
Qy 1801 TGTGTGGGACACAGCCAGTGGAGTGAAGTGTCTCACTCCCGGAGGATGGACTAGAG 1860
Db 1888 TGTGTGGGACACAGCCAGTGGAGTGAAGTGTCTCACTCCCGGAGGATGGACTAGAG 1947
Qy 1861 GTGTGTGTGACCCAGGCGCATGGGCGTTATGCTTGGCAGTGTCAAGAGGTTGGAGCC 1920
Db 1948 GTGTGTGTGACCCAGGCGCATGGGCGTTATGCTTGTGAATGTCAAGAGGTTGGGCA 2007
Qy 1921 GCCCGCTGTGTGTCTTATAGCTTGTGTGGGCGACGAGCGGGACCTCAAGCCGG 1980
Db 2008 GCCCATGTGTAGCAGCTTACAGCTTGGTATGGGCGACGACGAGATGCTCCAGCGG 2067
Qy 1981 GCCCACAGTGTGGGCGGTGGATGTGGTGGCTTTCTCCTGGTGTCTTTCAGCATCC 2040
Db 2068 GCCCACAC---AGTGGGGGGGAGCTGGCTGGCTTCTTGGGGATTCTCGCAGCATCC 2124
Qy 2041 CTCACCTCTCTCTGATTTGTCGCGCTCAGCAGCGTGGCGACAGAGGGAGCTTCTAGCT 2100
Db 2125 CTGACTCTCATCTTGTGTGGGCGTCAGCAGCGGCGGACAGAGGAACTTCTGGCT 2184
Qy 2101 AGAGACAAGTGGGCTTAGATCTGGGGCTCCACCTTCTGGGACACAAAGCTATAGTCAG 2160
Db 2185 AGAGACAAGTGGGCTGGACCTGGGCTCCACCTTCTGGGACACAAAGCTATAGTCAG 2244
Qy 2161 GACCTCTCTCTCTGCTGGAAGATGAAGCGTGGCCCTGGCGCTGGGTGAAGCGGGGC 2220
Db 2245 GACCTCTCTCTCTCTGGAAGATGAAGCGTGGCGCTGGCGCTGGGTGAAGCGGGGC 2304
Qy 2221 AGTGGTTTGTGTGCTTCCCTCCACCTTCTCTGCTGATCTTCCCAAGCCAGCCAC 2280
Db 2305 AGTGGTTTGTGTGATTTCTACACCTTCTCTGCTGATCTTCCCAAGCCAGCCAC 2364
Qy 2281 ATCCGGCTCACTGGGGCGCTCTAGCCACGCTGTGATGAGACCTCATCTA 2330
Db 2365 ATTCGGCTAACTGGGCTCTCTAGCCACATGTGATGAACATCCATCTA 2414

RESULT 6

BD123524
LOCUS BD123524 2971 bp DNA linear PAT 18-SEP-2002
DEFINITION Secretory protein or membrane protein.
ACCESSION BD123524
VERSION BD123524.1 GI:23218469
KEYWORDS JP 2002017376-A/33.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2971)
Ota,T., Isogai,T., Nishikawa,T., Kawai,Y., Sugiyama,T. and Hayashi,K.
REFERENCE 1
AUTHORS Secretory protein or membrane protein
TITLE Patent: JP 2002017376-A 33 22-JAN-2002;
JOURNAL HELIX RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
PN JP 2002017376-A/33
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253173
PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,YURI KAWAI,TOMOYASU
PI SUGIYAMA,
PI KOJI HAYASHI
PI PC

C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10, C12P21/02,C12O1/68/C12P21/08,C12N15/00,C12N5/00 CC
Secretory protein or membrane protein
FH Key Location/Qualifiers
FT CDS (103)..(2412).
FEATURES
source
1..2971
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 592 a 853 c 858 g 678 t
ORIGIN
Query Match 78.4%; Score 1827.6; DB 6; Length 2971;
Best Local Similarity 87.3%; Pred. No. 0;
Matches 2033; Conservative 0; Mismatches 279; Indels 18; Gaps 2;
Qy 1 ATGCTTCCAGGGCCGAGCCGCCCGCCCGGCGCCCGCGCGCTCCCGTTCCTTCCTTC 60
Db 103 ATGCCGGCTCTGCTGCGCGCCCGCCCGCGCTCCCGCGAGCTACAGCTTCGCCCTTC 162
Qy 61 CCGCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db 163 CC-----GCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 207
Qy 121 GTCCCGCGCTCAGTGGCCAGAACCTCGTGCCTCCATCTCCGAGGCTGACTCTCTATCTACC 180
Db 208 GTCCCGCGCTCGGTGCGCCAGAACCTCGTTCCTCAATCTCTGAGGCTGACTCTCTGCTACC 267
Qy 181 CGGTTTCCAGGCTTCATACATACATTAATTAATTAATTAATTAATTAATTAATTAATTA 240
Db 268 CGGTTTCCAGGCTTCATACATACATTAATTAATTAATTAATTAATTAATTAATTAATTA 327
Qy 241 ACACCTTACGTCGCTGCGGAGATAGCATCTTCGCTTTTAACCTTCCCTTCTCTGGGAA 300
Db 328 ACACCTTATGTTGGCGCCCGGACACCATCTTCCTCTTATCCCTGCTGCTGCTGCTGCTGCTG 387
Qy 301 AGACCCGAGGATCGACTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
Db 388 AGACCCGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 447
Qy 361 GGCAAGAAAGAGGAGCAATGTCAAAATTTATCAAGATTTCTGCGCATTTGTCATGCTCTCT 420
Db 448 GGCAAGAAAGAGGAGCAATGTCAAAATTTATCCAGATTTCTGCGCATTTGTCATGCTCTCT 507
Qy 421 CACCTCTCAGCTGCGGCGCTTCTGCTTTCATCCGAGTGGCGGTTATTTGATGCTGCTCC 480
Db 508 CACCTCTCAGCTTGTGGCACCCTTCGCTTTTATCCGAGTGGCGGTTATTTGATGCTGCTCC 567
Qy 481 AGTTTCCAGCAGGTTGAAAGACTTTCAGAGCGCGCGGGGAAATGTCTTTTTCAGCCAGCT 540
Db 568 AGTTTCCAGCAGGTTGAAAGACTTTCAGAGTGGCGCGGGGAAATGTCTTTTTCAGCCAGCT 627
Qy 541 CAACGGTCAGCAGTGTAAATGGTGGGCGCTCTCTCTACACCGGCGCTGTCGTCGTCGTCGTCG 600
Db 628 CAGCGTCAGCAGTGTAAATGGTGGGCGCTCTCTCTATGCTGCGCACTGTGAAACACTAC 687
Qy 601 CTGGGGAGTGAAGCCATCATCTCCGAGCTGCTGGTCCAGCTGAGGAGTGGATTCGGAACA 660
Db 688 CTGGGGAGTGAAGCCATCATCTCAGAGAGTGGTTCGTGGCGGAGGAGTGGATTCGGAACA 747
Qy 661 GAGACCTTGTCTATCTGCTTAAATGCTCCAGCTTTTTCGCGAGCTATGCTCTTCAGGCCA 720
Db 748 GATACCTTGTCTTCTGCTGCTGAGCCCGCCAGCTTTGTCGCGAGCGCTGCGCTTCGAGCCCA 807
Qy 721 GCTGAGTGGGGGATGAAGATGGAGCATGAATCTTTTTCCTTTCACCGAGACCTTC 780
Db 808 GCCAATGGGGGATGAAGATGGAGCATGAATCTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 867
Qy 781 CGAGTGTGGACTCTCTATGAGCGCATCAAGTCCCAAGAGTGGCCCGAGTCTGTCGGGG 840
Db 868 CGAGCATTTGACTCATAGAGCGCATTAAGTCCCAAGGTTGGCGGCTGTGTGTGTCGGGG 927

QY	841	GACCTTGGGGGAGGAGACCCCTTACAGAGAGATGGACGACCTTTCTGAAGCTGACCTG	900
Db	928	GACCTCGGGGGGGAAGACCCCTCCAGCAGAGATGGACGACGCTTTTGAAGCTGACCTG	987
QY	901	CTGTGCCAGGCGCCGAGCATGGCGGGGCTCGGGGTCTCGAGCTATGCGAGAGCTT	960
Db	988	CTCTGTCCAGGGCCTGAGCATGGCGGGGCTCCAGTCTCCTCAGGATGTTCTGTGCTT	1047
QY	961	CGGCTCAGCTGGAGCGGAACCCCATCTTTTATGGATCTTTTCCTCCAGTGGGAA	1020
Db	1048	CGACCTGAGCTTGGGCGAGGACTCCCATCTTTTATGGATCTTTTCCTCCAGTGGGAG	1107
QY	1021	GGAGTGCATCTCTGTGTGTGCTTCCGACCCCAAGACATCGGGCAGTGTGTAAT	1080
Db	1108	GGGGCTACTATCTCTGTGTGTGCTTCCGACCAAGACATTCGGACAGTGTGTAAT	1167
QY	1081	GGTCCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCCTGTCATGACACAGAG	1140
Db	1168	GGTCCCTTCAGAGAACTAAACATGACTGCAACAGAGGACTGCCTGTCGAGCAATGAT	1227
QY	1141	GTGCCCCAGCCAGACCTGGAGAGTGCATCGCAACAACATGAAGCTCCAGCAGTTTGA	1200
Db	1228	GTGCCCCAGCCAGACCTGGAGAGTGCATCACCACAACATGAAGCTCCGGCAGTTGGC	1287
QY	1201	TGCTCACTCTCCTTGCCAGACCGGTGCTACCTTTATCAGAGCACCCCTCTCATGGAC	1260
Db	1288	TCATCTCTCTCCTTGCCAGCGGTACTCAGCTTCATCCGGGAGCACCCCACTCATGGAC	1347
QY	1261	AGGCGCGTGTCCCGGCTACGCGCGCCCTGCTGCTCACTACAGATACGCCATCTC	1320
Db	1348	AGGCGAGTGTTCACAGCTGATGGCCACCCCTGCTGCTCACTACAGATACGCCATCTC	1407
QY	1321	AGAGTCTGTGGCCACAGGCTGACCAAGCTCTCAGGAAAGAAATATGACGTGCTACCTG	1380
Db	1408	AGAGTCTGTGGCCACAGGCTGACCAAGCTCTCAGGAAAGAAATATGACGTGCTACCTG	1467
QY	1381	GGGACAGAGATGGACACCTCCACCGGGGCTGTGGCGAATTTGGAGCTCAGTGTCTTG	1440
Db	1468	GGGACAGAGATGGACACCTCCACCGGAGCTGCGGATCGGAGCTCAGCTCAGCGTCTT	1527
QY	1441	GAGGATCTGCCTGTTCACAGACACAGCGGTTGAGAGCATGAAATGTACACGAT	1500
Db	1528	GAGATCTGGCCTTATCCAGAGCCACAGCAGTGTGAAGATGAAATGTACACAGC	1587
QY	1501	TGGCTCTGTGGGCTCCCATACTGAGGTGACACAAGTGAACACCAAGCACTGTGGCGCT	1560
Db	1588	TGGCTCTGTGGCTCCCGTACTGAGGTGACACAAGTGAATACACCAACTGTGGCGCT	1647
QY	1561	CTCCAGAGCTGTGGAGTGTATCTGTGGCCCAAGGACCCGCTGTGGCGCTGGAGCTTCGG	1620
Db	1648	CTCCAGAGCTGTCTAGAGTGCATCTCTGGCCCAAGGACCCAGTCTGTGGCTGGAGCTTC	1707
QY	1621	CTTGATGCTGTGTGCCCCACCGCGGAGCAGCCGGGATGGTTCACAGATACAGTCA	1680
Db	1708	CTGGATGAGTGTGTGGCCCATCCGCGGGAGCAGCGAGGTTGGTCAAGACATAGATCA	1767
QY	1681	GGGGATGCTCTCTTTGTGTCCAAAGAACCTGGAGAACATCCCGTAGTGTTTGAAGTT	1740
Db	1768	GCAGATGCTCTCTTTGTGTCTTAAGAGCCTGGAGAACGTCAGTAGTGTTTGAAGTT	1827
QY	1741	CCGGTGGCTACTGTGGGCGACGTGTCTGTCATGTTCCCGCAGTTCTGCCTGGGCATCC	1800
Db	1828	CCCGTGGCTACAGCTGCGCATGTGTCTTGGCCATGTTCTCCAAGCTCAGCATGGGCATCC	1887
QY	1801	TGTGTGTGSCACCCAGCCAGTGGAGTACTGGCTCACTCCCGGAGGAGTGGACTAGAG	1860
Db	1888	TGTGTGTGSCACCCAGCCAGTGGAGTACTGGCTCACTCACCCTCCCGGAGGAGTGGAG	1947
QY	1861	GTGGTGTGACCCAGGGGCACTGGGGGCTTATGCTTGGAGTGTGAGGAGGTGGAGCC	1920
Db	1948	GTGGTGTGACCCAGGGGCACTGGGGGCTTATGCTTGTGATGTGAGGAGGTGGGCA	2007
QY	1921	GCCCCGGTGGTGGCTGCTTATAGCTTGTGTGGGCGAGCCAGCGGGGACCCCTCAAAACCGG	1980
Db	2008	GCCCCATGTGGTAGCAGCTTACAGCTTGGTATGGGCGAGCCAGCAGATGCTCCGAGCCG	2067
QY	1981	GCCCCACCGTTGTGGGGCTGATTTGGTTGGCTTCTCCTGGTCTTCTTCAGCAGCATCC	2040
Db	2068	GCCCCACAC---AGTGGGGCGGAGCTGGCTGCTTCTTCTGGGATTCGCAAGCATCC	2124
QY	2041	CTCACTCTCTCTGANTTGGTCCGCTCAGCAGCGTGGCGACAGAGGGAGCTTCTAGCT	2100
Db	2125	CTGACTCTCATTTCTGTTGGCTCAGAGCGAGCGCAGACAGAGGAACTTCTGGCT	2184
QY	2101	AGAGACAAGTGGGCTTAGATCTGGGGGCTCCACCTCTTGGGACCAACAAGCTATAGTCA	2160
Db	2185	AGAGACAAGTGGGCTGGACCTGGGGGCTCCACCTTCTGGGACCAACAAGCTACAGCAA	2244
QY	2161	GACCTCTCCCTCTCGCTCGCCTGAAGATGAACGCTGCCCTGGCCCTGGCTGAAGCGGGC	2220
Db	2245	GACCTCTCCCTCCCTCTCTCTGAGATGAGCGGTGCCCCTGGCCCTGCCAAGAGGGGC	2304
QY	2221	AGTGGTTTTTGGTGGCTTCCCTCCACCTTCTCTGCTGGATTTGGCCCAAGCCAGCCAC	2280
Db	2305	AGTGGCTTTGGTGGATTTCTCACACCTTCTCTGCTTGATCTTGGCCCAAGCCAGCCAC	2364
QY	2281	ATCCGGCTCACTGGGGGCTCTAGCCAGCTGTGATGAGACCTCCATCTA	2330
Db	2365	ATTCGGCTAACTGGGGCTCTCTAGCCACATGTGATGAACATCCATCTA	2414
RESULT 7			
AK075384		2971 bp mRNA linear PRI 03-SEP-2002	
LOCUS		Homo sapiens cDNA PSEC0074 fis, clone NT2RP2003050, highly similar	
DEFINITION		to Mus musculus mRNA for semaphorin W.	
ACCESSION		AK075384	
VERSION		AK075384.1 GI:22761435	
KEYWORDS		oligo capping; fis (full insert sequence).	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS		Ota,T., Nishikawa,T., Suzuki,Y., Kawai-Hio,Y., Kawai-Hio,Y., Hayashi,K., Ishii,S., Saiko,K., Yamamoto,J., Wakamatsu,A., Nagai,T., Nakamura,Y., Nagahari,K., Sugano,S. and Isoigai,T.	
TITLE		HRI human cDNA sequencing project	
JOURNAL		Unpublished	
REFERENCE		2 (bases 1 to 2971)	
AUTHORS		Isoigai,T. and Yamamoto,J.	
TITLE		Direct Submission	
JOURNAL		Submitted (20-MAR-2002) Takao Isoigai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-53-3975, Fax:81-438-53-3986)	
COMMENT		HRI human cDNA sequencing project; cDNA 5' & 3'-end one pass sequencing, clone selection and full insert sequencing; Helix Research Institute (supported by Japan Key Technology Center etc.); cDNA library construction; Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center.	
FEATURES		Location/Qualifiers	
source		1..2971	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="NT2RP2003050"	
		/cell_line="NT2"	
		/cell_type="teratocarcinoma"	
		/clone_lib="NT2RP2"	
		/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction"	
		103..2415	
		/note="unnamed protein product"	
		/codon_start=1	
CDS			


```
Qy 601 CTGGGGACTGAGCCCATCATCTCCCGAGCTGTGGGTGAGCTGAGGACTGATTCGAACA 660
Db 697 CTGGGACGGAGCAATTATCACCAGAGCAGTGGTGGTCCGAGGACTGATTTCGGACA 756
Qy 661 GAGACCTTGTCATCTCGCTTAATGCTCCAGCCCTTTCTCGCAGCTATGTTCTCAGCCCA 720
Db 757 GATACCTTGCTCTCTGGCTGGAAGCCGCCACCCCTTTCTCGCAGCCGTTGGCTTGAGCCCA 816
Qy 721 GCTGAGTGGGGGATGAAGATGGAGACGATGAATCTTTTCTTCTTTCACGGAGACCTCC 780
Db 817 GCCGAATGGGGATGAAGATGGAGACGACGAATCTACTTCTTCTTACGGAGACTTCC 876
Qy 781 CGAGTGTGGACTCTATGAGCCGATCAAGTCCCAAGATGGCCCCAGTGTGTGCGGG 840
Db 877 CGAGCATTTGACTCATACGAGCGCATTAAGATCCCAAGTGGCCCGCTGTGTGCGGG 936
Qy 841 GACCTTGGGGGAGGAGACCCCTTCACGAGAGATGGACGACTTTCCTGAAGGCTGACCTG 900
Db 937 GACCTTGGGGGCGGAAGACCCCTCCACGAGAGATGGACGAGCTTTTGAAGCTGACCTG 996
Qy 901 CTGTGCCAGGGCCGAGCATGGCCGGGCTTCGCGGGTCTCTGAGGCTATGGCAGAGCTT 960
Db 997 CTCTGTCCAGGGCTGAGCATGGCCGGCTCCAGTGTCTCTGAGGATGTTGCTGTCTT 1056
Qy 961 CGGCTCAGCTGGAGCGGGAACCCCATCTTTTATGGGATCTTTTCCCTCCAGTGGGAA 1020
Db 1057 CGACCTGAGCTTGGGGAGGAGACTCCCATCTTTTATGGCATCTTTTCCCAAGTGGGAG 1116
Qy 1021 GGAGTCCCATCTCTGTGTGTGCTTCGACCCCAAGACATCCGGGAGTGTCTCAAT 1080
Db 1117 GGGCTACTATCTGTCTGTCTGCTTCCGACCAAGACATTCGGCAGTGTCTGAAT 1176
Qy 1081 GGTCCCTTTAGAGAGCTTAAACATGACTGCAACAGGGGACTGCTGTCTATGGACAACGAG 1140
Db 1177 GTTCCCTTCAGAGACTTAAACATGACTGCAACAGAGGACTGCTGTCTGGCAATGAT 1236
Qy 1141 GTGCCCCAGCCAGACCTCGAGAGTGCATGCCCAACAAATGAAGCTCCAGAGTTGGA 1200
Db 1237 GTGCCCCAGCCAGACCTCGAGAGTGCATCAACAAACATGAAGCTCCGCGCACTTGGC 1296
Qy 1201 TCCTACTCTCCCTGCCAGACCGGTCTCACTTTTATCAGAGACCAACCTCTCATGAC 1260
Db 1297 TCATCTCTCCCTGACCGGTACTCACTTTCATCCGGGACCAACCTCATGGAC 1356
Qy 1261 AGCCCCGTGTTCCCGGTGACGCCGCCCTGCTGTCTACTACAGATACAGCCTATCTC 1320
Db 1357 AGGCCAGTGTTCAGCTGATGGCCACCCCTGCTGTCTACTACAGATACAGCCTATCTC 1416
Qy 1321 AGAGTGTGCCCCACAGGTGACAGCCTCTCAGGGAAGAAATATGACGTGCTCTACCTG 1380
Db 1417 AGAGTGTGCCCCACAGGTGACAGCCTCTCAGGGAAGAGTATGATGTGCTCTACCTG 1476
Qy 1381 GGGACAGAGATGACACCTCCACCGGCTGTGGCATTTGAGCTCAGCTCAGTGTCTTG 1440
Db 1477 GGGACAGAGATGGACACTCCACCGAGAGTGGGATCGGAGCTCAGCTCAGCGTTCTT 1536
Qy 1441 GAGGATCGCTTGTGTTCCAGAACACAGCCGGTTGAGAGCATGAATTTACACAGAT 1500
Db 1537 GAAGATCGCTTATTTCCAGAGCCACAGCCAGTTCAGAACATGAATTTACACACAGC 1596
Qy 1501 TGGCTCTGTGGCTCCCATACTAGGTGACACAGTGAACACCACTGTGGCCGT 1560
Db 1597 TGGCTCTGTGGTCTCCCTACTAGGTGACACAGTGAATACAACCACTGTGGCCGT 1656
Qy 1561 CTCAGAGCTGCTCGGAGTGTATCTCTGCCCCAGACCCCGTGTGCGCTTGAGCTTCCGG 1620
Db 1657 CTCAGAGCTGCTCAGAGTGCATCTGTGGCCAGAGCCAGTGTGTGCTGGAGCTTCCGG 1716
Qy 1621 CTTGATCTGTGGCCCCAGCGGGGAGCACCGCGGATGTTCAAGATATAGATCA 1680
Db 1717 CTGATCAGTGTGTGGCCCATCGCGGGAGACCGAGGTTGTGTCGAAGACATAGATCA 1776
Qy 1681 GCGGATCTCTCTTCTTGTGTCCAAAAGAACCTGGAGAACATCCCGTAGTGTGTTGAAGTT 1740
```

```
Db 1777 GCAGATCTCTCCTTTTGTGTCTTAAAGAGCCTGGAGAACCTCCAGTAGTGTGTTGAAGTT 1836
Qy 1741 CCGTGTGCTACTGTGGGCCAGTGGTCTCGCATGTTCCCCAGTTCCTGCTGGCATCC 1800
Db 1837 CCGTGTGCTACAGTGGCATGTGGTCTTGCCATGTTCTTCCAACTCAGCATGGCATCC 1896
Qy 1801 TGTGTGTGGACCAAGCCAGTGGAGTGCCTGCTCCTACTCCCCGAGGAGTGGAGTACAG 1860
Db 1897 TGTGTGTGGACCAAGCCAGTGGAGTGCCTGCTCCTACTCCCCGCGGGATGGAGTGGAG 1956
Qy 1861 GTGTGTGTGACCCAGGGGCCATGGGGCTTATGCTTTCGAGTGTCTAGGAGGTTGGAGCC 1920
Db 1957 GTGTGTGTGACCCAGGGGCCATGGGGCTTATGCTTCTGAATGTCTAGGAGGTTGGGCA 2016
Qy 1921 GCCGCGCTGTGGCTTATAGCTTGTGTGGGAGCAGCAGCGGGACCTCAAAACGG 1980
Db 2017 GCCCATGTGTAGCAGCTTACAGCTTGTATGGGGCAGCAGCAGATGCTCCGAGCGG 2076
Qy 1981 GCCACACCGTGTGGGGCTGGATTGTTGGCTTTCTCTCTGGTGTGTTTGCAGCATCC 2040
Db 2077 GCCACACAC---AGTGGGGCGGAGTGGCTGGCTTCTTGGGGATTCTCGCAGCATCC 2133
Qy 2041 CTCACTCTCTCTGATTGTGCGCGTTCAGCAGCGTTCGGGACAGAGAGGAGCTTCTAGCT 2100
Db 2134 CTGACTCTCATCTTGTATTGTTGGCGTTCAGCAGCGAGCGGACAGAGGAACTTCTGGCT 2193
Qy 2101 AGACACAAGTGGGCTTAGATCTGGGGCTCCACCTTCTGGGACCAACAGTATAGTACAG 2160
Db 2194 AGACACAAGTGGGCTTGGACCTGGGGCTCCACCTTCTGGGACCAACAGTATAGCACA 2253
Qy 2161 GACCTCTCTCTCTTCCCTGAAAGATGAACGGCTGCGCTGGCCCTGGCTTAAGCGGGGC 2220
Db 2254 GACCTCTCTCTCTCTCTCTGAAAGATGAGCGTTGCGCTGGCCCTGGCCCAAGAGGGC 2313
Qy 2221 AGTGGTTTGTGTGCTTCCCTCCACCTTCTCTGCTGCTGATGTTCTGCCCAAGCCAGCCAC 2280
Db 2314 AGTGGCTTGTGTGATTTCTCACCACTTCTCTGCTGCTGATCTTGTCCCAAGCCAGCCAC 2373
Qy 2281 ATCCGGCTCACTGGGCGCCTCTAGCCACGTTGATGAGACTCCATCTA 2330
Db 2374 ATTGGCTACTGGGGCTCTCTAGCCACATGTCATGAACATCCATCTA 2423
```

```
RESULT 9
AX003081 LOCUS AX003081 2893 bp DNA linear PAT 24-AUG-2000
DEFINITION Sequence 1 from Patent EP0933425.
ACCESSION AX003081
VERSION AX003081.1 GI:9926964
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Michalovich,D. and Doe,T.R.
TITLE Semaphorin family polypeptides and polynucleotides
JOURNAL Patent: EP 0933425-A 1 04-AUG-1999;
SMITHKLINE BEECHAM PLC (GB)
FEATURES
source
1..2893
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 579 a 825 c 842 g 647 t
ORIGIN
```

```
Query Match 69.8%; Score 1626; DB 6; Length 2893;
Best Local Similarity 83.28; Pred. No. 0;
Matches 1938; Conservative 0; Mismatches 275; Indels 117; Gaps 3;
Qy 1 ATGCTTGCACGGGCGGAGCGGCCCGCGCGCGCGCGCGCTTTCCTTCCTTC 60
```


Qy	241	ACACTTTACGTCGGTGCACGGGATACGACTCTTCGCTTTTAAACCTCCCTCTCTCGGGAA	300
Db	335	ACACTTTATGTTGGCGCCGGACACCATCTTCGCTTTATCCCTGCCCCCTCTCAGGGAG	394
Qy	301	AGACCCGAAGGATCGACTGGATGGTACCTGAGACTCACAGACAGAACTGCAGGAAGAA	360
Db	395	AGACCCGCAGGATGACTGGATGGTTCCTGAGGCTCACAGACAGAACTGTAGAGAGAA	454
Qy	361	GGCAAGAAAGAGGACGAATGTCACAAATTTTATCCAGATTCCTGCCATGTGTCAATGCCTCT	420
Db	455	GGCAAGAAAGAG-----	466
Qy	421	CACCTCCTCAGCTGGCGACCTTCGCTTTTGTATCCGAGTGGCGGGTTATTGATGTGTCC	480
Db	467	-----GATGTGTCC	475
Qy	481	AGTTTCAGCAGGTTGAAGACTTGAGAGCGCGGGGGAATATGCTCTTTTGAGCCAGCT	540
Db	476	AGGTTCCAGAGGTTGAAGACTTGAGAGTGGCGGGGGAATATGCTCTTTTGAGCCAGCT	535
Qy	541	CAACGGTTCAGCAGCTGTAAATGGCTGGGGCGCTCCTCTACACCCGCCACTGTGAAGAATTC	600
Db	536	CAGCGGTTCAGCAGCTGTAAATGGCTGGGGGGTCTCTATGCTGCCACTGTGAAAACTAC	595
Qy	601	CTGGGNACTCAGCCCATCATCTCCGAGCTGTGGGTGCGAGCTCAGAGCTGGATTCGNACA	660
Db	596	CTGGGACGGAGGCAATATCACAGAGCAGTGGGTCTGCCGAGGACTGGATTCGGACA	655
Qy	661	GAGACCTTGTCATCCTGGCTTAATGCTTCACAGCCTTTGTCGACGCTATGGTCCTGAGGCCA	720
Db	656	GATACCTTGCTTCCTGGCTGACGCCCGAGCCTTTGTCTGACGCGTGGCCTTGAGCCCA	715
Qy	721	GCTGAGTGGGGGATGAAGATGGAGACGATGAATCTTTTTTTTCTTCACGGAGACCTCC	780
Db	716	GCCGAATGGGGGATGAAGATGGAGACGACGAAATCTACTTCTTTACGGAGACTTCC	775
Qy	781	CGAGTGTGACATCTATGAGCGCATCAAGGTCCCAAGATGGCGCGAGTGTGTGCGGG	840
Db	776	CGAGCATTTGACTCATACGACGCAATTAAGTCCCGGGTGGCGCGCTGTGTGTGCGGG	835
Qy	841	GACCTTGGGGCAGGAAGACCTTCACGACAGATGGACGACTTCTCTGAAGGCTGACCTG	900
Db	836	GACCTCGGGGGCGGAAGACCTTCACGACAGATGGACGACTTTTGAAGCTGACCTG	895
Qy	901	CTGTGCCAGGGGCCGAGCATGSCCGGGCCTCCGGGTTCTCGAGGCTATGGCAGAGCTT	960
Db	896	CTCTGTCCAGGGCTTGACATGSCCGGGCCTCCAGTCTCTCGCAGGATGTGTCTGTCTT	955
Qy	961	CGGCTCAGCCTCGAGGGGAAACCCCATCTTTTATGGATCTTTTCTCCCAGTGGGAA	1020
Db	956	CGACTGAGCTTGGGCAGGGAGCTCCCATCTTTTATGGCATCTTTCTCCAGTGGGAG	1015
Qy	1021	GGAGCTCCCATCTCTGCTGTGTGCTTCCGACCCCAAGACATCCGGGAGTGCCTGAAT	1080
Db	1016	GGGGCTACTATCTCTGCTGTCTGTGGCTTCCGACCACAAGACATTCGGCAGTGTCTGAAT	1075
Qy	1081	GGTCCCTTTAGAGAGCTAAAACATGACTGCAACAGGGGACTGCTGTCTATGGACAACGAG	1140
Db	1076	GGTCCCTTCAGAGACTAAAACATGACTGCAACAGAGGACTGGCTGTCTGTGGACAATGAT	1135
Qy	1141	GTGCCCGACCCAGACCTGGAGAGTGCATCGCCCAACAACTAAGCTCCAGCAGTTTGA	1200
Db	1136	GTGCCCGACCCAGACCTGGAGAGTGCATCACCACAACATGAAGCTCCGGCACTTTGGC	1195
Qy	1201	TCCTCACTCTCCCTGCCAGACCGGGTCTCACCTTTATCAGAGACCAACCTCTCATGAGAC	1260
Db	1196	TCATCTCTCTCCCTGCTGACCGCGTACTCACCTTATCCGGGACCAACCACTCATGGAC	1255
Qy	1261	AGGCCGTGTTCCGGCTGACGGCGCCCTGCTGGTGTCACTACAGATACAGCCTATCTC	1320
Db	1256	AGGCCAGTGTTCAGCTGTAGCCACCCCTGCTGTGTCACTACAGATACAGCCTATCTC	1315
Qy	1321	AGAGTCGTGGCCACAGGGTGACACGCTCTCAGGGAAGAATATACGTGTCTACCTG	1380

[illegible][illegible]

RESULT 11
E35443
LOCUS
DEFINITION
ACCESSION

[illegible]

Best Local Similarity 83.2%; Pred. No. 0; Matches 1938; Conservative 0; Mismatches 275; Indels 117; Gaps 3;									
Qy	1	ATGTCCTCCAGGCGGAGCGGCCCGCCGCGCGCCCGCGCGCTCTCCGGTCTTTCCCTTC	60						
Db	110	ATGCGGCTCTGCTGCGCGCGCCGCGCGCGCGCGCGAGCTCAGAGCTCGCCCTTC	169						
Qy	61	CCGCGCGCGTGTGCTGCTGCTGCTGCGGATATAAGCGCGCCGGTGTGCGCGCG	120						
Db	170	CC-----GCTACTGCTGCTGGGCTGCTGAGCGCCGGTATCGGCGCG	214						
Qy	121	GTCCCCGCTCAGTGCCCAAGACCTCGCTGCCATCTCCGAGGCTGACTCCTATCTCAAC	180						
Db	215	GTCCCCGCTCGGTGCGCCAGAACCTGCTTCCRAATCTGAGGCTGACTCCTGTCTCACC	274						
Qy	181	CGTTTTACAGCGTCTATACGTACAAATTACTCTCTCTCTCTGTTGGAATCTGCTCCAC	240						
Db	275	CGTTCGCACTCCCTACACATACAAATTACTGTCTCTCTGTTGGAATCTGCTCCAC	334						
Qy	241	ACACTTTACGTGCGGTGCACGGATAGCATCTTCGCTTTAAACCTTCCCTTCTCTGGGAA	300						
Db	335	ACACTTTATGTGGCGCGCGGACACCATCTTGCTTTATCCCTGCCCCTTCTCAGGGGAG	394						
Qy	301	AGACCCGAGGATCGACTGGATGGTACCTGAGACTCACAGACAGAAGTGCAGGAAGAAA	360						
Db	395	AGACCCGCGCAGGATGTGATGGATGGTTCCTGAGGCTCACAGACAGAAGTGTAGGAAGAAA	454						
Qy	361	GGCAAGAAAGAGGACGAATGTCACAATTTATCCAGATTCTCGGCAATTGTCAATGCCTCT	420						
Db	455	GGCAGAAAGAG-----	466						
Qy	421	CACCTCTCAGTGCAGCACCTTGCTTTTATCCGAAGTGCGGGTTATTGATGTCTCC	480						
Db	467	-----GATGTCTCC	475						
Qy	481	AGTTTCAGCAGGTTGAAAGACTTGAGAGCGCGCGGGGAAATGTCCTTTTGAGCCAGCT	540						
Db	476	AGGTTCCAGCAGGTTGAAAGACTTGAGAGTGGCGGGGAAATGTCCTTTTGAGCCAGCT	535						
Qy	541	CAACGGTTCAGCAGCTGTAATGGCTGGGGGGTCTCTACACCGGCACCTGTGAAGAACTTC	600						
Db	536	CAGCGTTCAGCAGCTGTAAATGGCTGGGGGGTCTCTATGCTGCCATGTGAAAAACTAC	595						
Qy	601	CTGGGAGCTAGCCCATCATCTCCGAGCTGTGGGTCGAGCTGAGGACTGGATTCGAACA	660						
Db	596	CTGGGAGCGAGCCAAATATCACAGACAGTGGTGTGTCGCGAGGACTGGATTCGACA	655						
Qy	661	GAGACCTTGATCCTGGCTTAATGCTCCAGCCTTTGTCGAGGTATGTCCTGAGCCCA	720						
Db	656	GATACCTTGGCTTCTCTGGCTGAACGCCCCAGCCTTTGTCGAGCGCGTGGCTTGAGCCCA	715						
Qy	721	GCTCAGTGGGGGATGAAGATGGAGACGATCAATCTTTTCTTTCTTCGAGACCTCC	780						
Db	716	GCCGAATGGGGGATGAAGATGGAGACGAAATCTACTTCTTTTTCGAGAGACTTCC	775						
Qy	781	CGAGTGTGGACTCCTATGACGCGATCAAGTCCCAAGATGGCCCGAGTGTGTGCGGG	840						
Db	776	CGACATTTGACTCATACGAGCGCATTTAAAGTCCACGGGTGGCCCGTGTGTGTCGGGG	835						
Qy	841	GACCTTGGGGCAGGAAGACCTTTCAGCAGATGGACAGCTTTTGAAGGCTGACCTG	900						
Db	836	GACCTCGGGGCGGAAGACCTCCAGCAGAGATGGACAGCTTTTGAAGGCTGACCTG	895						
Qy	901	CTGTGCCCGCGCGAGATGGCCGGGCTCGGGGTTCTGACGAGCTATGGCAGAGCTT	960						
Db	896	CTCTGTCCAGGGGCTGAGCATGCGCGGGCTCCAGTGTCTGCAAGATGTGTCTGTCTT	955						
Qy	961	CGGCTCAGCTGGAGCGGAACCCCATCTTTTATGGGATCTTTTCTCCAGTGGAA	1020						
Db	956	CGAGCTGAGCTTGGGCGAGGACTCCCATCTTTTATGCGATCTTTTCTCCAGTGGAG	1015						
Qy	1021	GGAGCTGCCATCTCTGCTGTGTGTGCTTCCGACCCCAAGACATTCGCGGAGTGTCTGAAT	1080						

Db 1548 CCCGTGGCTACAGTGGCCATGTGGTCTTGCCATGTTCTCAAGCTCAGCATGGGCATCC 1607
QY 1801 TGTGTGTGGCACCAGCCAGTGGAGTGAAGTCTGCTCACTCCCGGAGGGATGGACTAGAG 1860
Db 1608 TGTGTGTGGCACCAGCCAGTGGAGTGAAGTCTGCTCACTCCCGGAGGGATGGACTGGAG 1667
QY 1861 GTGGTGTGACCCACAGGGGCCATGGGGGCTTATGCTTCGAGAGTGTACAGAGGTGGAGCC 1920
Db 1668 GTGGTGTGACCCACAGGGGCCATGGGGGCTTATGCTTCGAGAGTGTACAGAGGTGGAGCC 1727
QY 1921 GCCCGCTGGTGGCTGTATAGCTTGGTGGGCGACCCAGCGGGGACCTCAACCCG 1980
Db 1728 GCCCATGGGTAGCAGCTTAGAGCTTGGTGGGCGACCCAGAGATGCTCCGAGCCGG 1787
QY 1981 GCCCACACGTTGTGGGGCTGGATTTGGCTTTCTCTGGGTGTTCTTCAGCATCC 2040
Db 1788 GCCCACAC---AGTGGGGCGGGACTGGCTGGCTTCTTTGGGATTTCTGCAGCATCC 1844
QY 2041 CTCAGTCTCTCTGATTGGTGGCGGTGAGCGGTGAGCGGTGCGGACAGAGGGAGTCTTAGCT 2100
Db 1845 CTGACTCTCATTTCTGATTGGTGGCGGTGAGCGGTGAGCGGTGCGGACAGAGGGAATCTTGCT 1904
QY 2101 AGAGACAAGGTGGGCTTAGATCTGGGGGCTCCACCTTCTGGGACCACAAAGCTATAGTCAG 2160
Db 1905 AGAGACAAGGTGGGCTGGACCTGGGGGCTCCACCTTCTGGGACCACAAAGCTACAGCCAA 1964
QY 2161 GACCTCCCTCTCCTTCGCTGAGATGAAGCGGTGCGCCCTGGCCCTGGTAAAGCGGGC 2220
Db 1965 GACCTCCCTCCCTCTCCTGAGATGAGCGGTGCGCCCTGGCCCTGGCCCAAGAGGGC 2024
QY 2221 AGTGGTGTGGTGGCTTCCCTCCACCCCTTCTGCTGATTTCTTGCCCAAGCCCAAGCCAC 2280
Db 2025 AGTGGCTTTGTGGATTCTCACCACCCCTTCTGCTGATCCTTGCCCAAGCCCAAGCCAC 2084
QY 2281 ATCCGGCTCACTGGGGGCTCTAGCCAGCTGTGATGAGACCTCCATCTA 2330
Db 2085 ATCCGGCTCACTGGGGGCTCTAGCCACATGTGATGAACATCCATCTA 2134

Search completed: September 26, 2003, 02:11:47
Job time : 8353.28 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2003, 19:27:00 ; Search time 11674.4 Seconds
(without alignments)
11385.680 Million cell updates/sec

Title: US-09-284-180A-1
Perfect score: 4008
Sequence: 1 gccgagccgcgcagtagc.....aaaaaaaaaaaaaaaaaaaa 4008

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 33363688 seqs, 1658189874 residues

Total number of hits satisfying chosen parameters: 66727376

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main.*
1: /cgn2_6/ptodata/2/pna/pctus_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/pctus_COMB.seq.old.*
3: /cgn2_6/ptodata/2/pna/US06_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US07_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US080_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US081_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US083_COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US084_COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US085_COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US086_COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US088_COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US089_COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US090_COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US091_COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US092A_COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US092B_COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US093A_COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US093B_COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US094_COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US095A_COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US095B_COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US095C_COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US095D_COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US096A_COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US096B_COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US096C_COMB.seq.*
29: /cgn2_6/ptodata/2/pna/US096D_COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US096E_COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US097A_COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US097B_COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US097C_COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US098A_COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US098B_COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US098C_COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US098D_COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US099A_COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US099B_COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US099C_COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US099D_COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US099E_COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US099F_COMB.seq.*

44: /cgn2_6/ptodata/2/pna/US100A_COMB.seq.*
45: /cgn2_6/ptodata/2/pna/US100B_COMB.seq.*
46: /cgn2_6/ptodata/2/pna/US101A_COMB.seq.*
47: /cgn2_6/ptodata/2/pna/US101B_COMB.seq.*
48: /cgn2_6/ptodata/2/pna/US102A_COMB.seq.*
49: /cgn2_6/ptodata/2/pna/US102B_COMB.seq.*
50: /cgn2_6/ptodata/2/pna/US103A_COMB.seq.*
51: /cgn2_6/ptodata/2/pna/US103B_COMB.seq.*
52: /cgn2_6/ptodata/2/pna/US104A_COMB.seq.*
53: /cgn2_6/ptodata/2/pna/US104B_COMB.seq.*
54: /cgn2_6/ptodata/2/pna/US6000_COMB.seq.*
55: /cgn2_6/ptodata/2/pna/US6001_COMB.seq.*
56: /cgn2_6/ptodata/2/pna/US6002_COMB.seq.*
57: /cgn2_6/ptodata/2/pna/US6003_COMB.seq.*
58: /cgn2_6/ptodata/2/pna/US6004_COMB.seq.*
59: /cgn2_6/ptodata/2/pna/US6005_COMB.seq.*
60: /cgn2_6/ptodata/2/pna/US6006_COMB.seq.*
61: /cgn2_6/ptodata/2/pna/US6007_COMB.seq.*
62: /cgn2_6/ptodata/2/pna/US6008_COMB.seq.*
63: /cgn2_6/ptodata/2/pna/US6009_COMB.seq.*
64: /cgn2_6/ptodata/2/pna/US6010_COMB.seq.*
65: /cgn2_6/ptodata/2/pna/US6011_COMB.seq.*
66: /cgn2_6/ptodata/2/pna/US6012_COMB.seq.*
67: /cgn2_6/ptodata/2/pna/US6013_COMB.seq.*
68: /cgn2_6/ptodata/2/pna/US6014_COMB.seq.*
69: /cgn2_6/ptodata/2/pna/US6015_COMB.seq.*
70: /cgn2_6/ptodata/2/pna/US6016_COMB.seq.*
71: /cgn2_6/ptodata/2/pna/US6017_COMB.seq.*
72: /cgn2_6/ptodata/2/pna/US6018_COMB.seq.*
73: /cgn2_6/ptodata/2/pna/US6019_COMB.seq.*
74: /cgn2_6/ptodata/2/pna/US6020_COMB.seq.*
75: /cgn2_6/ptodata/2/pna/US6021_COMB.seq.*
76: /cgn2_6/ptodata/2/pna/US6022_COMB.seq.*
77: /cgn2_6/ptodata/2/pna/US6023A_COMB.seq.*
78: /cgn2_6/ptodata/2/pna/US6023B_COMB.seq.*
79: /cgn2_6/ptodata/2/pna/US6024_COMB.seq.*
80: /cgn2_6/ptodata/2/pna/US6025_COMB.seq.*
81: /cgn2_6/ptodata/2/pna/US6026_COMB.seq.*
82: /cgn2_6/ptodata/2/pna/US6027_COMB.seq.*
83: /cgn2_6/ptodata/2/pna/US6028_COMB.seq.*
84: /cgn2_6/ptodata/2/pna/US6029_COMB.seq.*
85: /cgn2_6/ptodata/2/pna/US6030_COMB.seq.*
86: /cgn2_6/ptodata/2/pna/US6031_COMB.seq.*
87: /cgn2_6/ptodata/2/pna/US6032_COMB.seq.*
88: /cgn2_6/ptodata/2/pna/US6033_COMB.seq.*
89: /cgn2_6/ptodata/2/pna/US6034_COMB.seq.*
90: /cgn2_6/ptodata/2/pna/US6035_COMB.seq.*
91: /cgn2_6/ptodata/2/pna/US6036_COMB.seq.*
92: /cgn2_6/ptodata/2/pna/US6037_COMB.seq.*
93: /cgn2_6/ptodata/2/pna/US6038_COMB.seq.*
94: /cgn2_6/ptodata/2/pna/US6039_COMB.seq.*
95: /cgn2_6/ptodata/2/pna/US6040_COMB.seq.*
96: /cgn2_6/ptodata/2/pna/US6041_COMB.seq.*
97: /cgn2_6/ptodata/2/pna/US6042_COMB.seq.*
98: /cgn2_6/ptodata/2/pna/US6043_COMB.seq.*
99: /cgn2_6/ptodata/2/pna/US6044_COMB.seq.*
100: /cgn2_6/ptodata/2/pna/US6045_COMB.seq.*
101: /cgn2_6/ptodata/2/pna/US6046_COMB.seq.*
102: /cgn2_6/ptodata/2/pna/US6047_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	4008	100.0	4008	18	US-09-284-180-1 Sequence 1, Appl
2	4008	100.0	4008	18	US-09-284-180A-1 Sequence 1, Appl
3	2331	58.2	2331	18	US-09-284-180-2 Sequence 2, Appl
4	2331	58.2	2331	18	US-09-284-180A-2 Sequence 2, Appl

Qy	1021	GCTATGCCAGAGCTTCGGCCTCAGCCTCGAGCGGGAACCCCAATCTTTTATGGGATCTTT	1080
Db	1021		
Qy	1021	GCTATGGCAGAGCTTCGGCCTCAGCCTGGAGCGGGAACCCCAATCTTTTATGGGATCTTT	1080
Db	1021		
Qy	1081	TCCTCCAGTGGGAAGAGCTGCATCTCTGCTGTGTGCTGCTTCGAGCCCAAGACATC	1140
Db	1081		
Qy	1081	TCCTCCAGTGGGAAGAGCTGCCATCTCTGCTGTGTGCTTCCGACCCCAAGACATC	1140
Db	1081		
Qy	1141	CGGGCAGTGTGAATGGTCCCTTTAGAGAGCTAAAAACATGACTGCAACAGGGGACTGCCT	1200
Db	1141		
Qy	1141	CGGGCAGTGTGAATGGTCCCTTTAGAGAGCTAAAAACATGACTGCAACAGGGGACTGCCT	1200
Db	1141		
Qy	1201	GTCAATGGACAACAGAGTGCACAGCCAGACCTGGAGAGTGCATCGCCAACACATGAAG	1260
Db	1201		
Qy	1201	GTCAATGGACAACAGAGTGCACAGCCAGACCTGGAGAGTGCATCGCCAACACATGAAG	1260
Db	1201		
Qy	1261	CTCCAGCAGTTTGGATCTCACTCTCCCTGCCAGACCGGTGCTCACTTTATCAGAGAC	1320
Db	1261		
Qy	1261	CTCCAGCAGTTTGGATCTCACTCTCCCTGCCAGACCGGTGCTCACTTTATCAGAGAC	1320
Db	1261		
Qy	1321	CACCTCTCATGACAGGCCGTGTTCCGGCTGACGGCGCCCTCTGTGTGTCATACTACA	1380
Db	1321		
Qy	1321	CACCTCTCATGACAGGCCGTGTTCCGGCTGACGGCGCCCTCTGTGTGTCATACTACA	1380
Db	1321		
Qy	1381	GATACAGCTTATCTCAGAGTCGTGGCCACAGGGTGACCAGCCTCTCAGGAAAGAATAT	1440
Db	1381		
Qy	1381	GATACAGCTTATCTCAGAGTCGTGGCCACAGGGTGACCAGCCTCTCAGGAAAGAATAT	1440
Db	1381		
Qy	1441	GACGTCTCTACCTGGGGACAGAGTAGACACTCCACCGGGCTGTGGGCATTTGGAGCT	1500
Db	1441		
Qy	1441	GACGTCTCTACCTGGGGACAGAGTAGACACTCCACCGGGCTGTGGGCATTTGGAGCT	1500
Db	1441		
Qy	1501	CAGCTCAGTGTCTTGAGAGTCTGGCCTTGTTCCTCAGAACACAGCCGGTTGAGAGCATG	1560
Db	1501		
Qy	1501	CAGCTCAGTGTCTTGAGAGTCTGGCCTTGTTCCTCAGAACACAGCCGGTTGAGAGCATG	1560
Db	1501		
Qy	1561	AAATTGTACCACGATTTGGCTTCTGTGGGTCCCTACTGAGGTGACACAAGTGAACACC	1620
Db	1561		
Qy	1561	AAATTGTACCACGATTTGGCTTCTGTGGGTCCCTACTGAGGTGACACAAGTGAACACC	1620
Db	1561		
Qy	1621	AGCAATGTGGCGTCTCCAGAGCTGCTCGGAGTGTACTCTGGCCACAGACCCCGTGTGC	1680
Db	1621		
Qy	1621	AGCAATGTGGCGTCTCCAGAGCTGCTCGGAGTGTACTCTGGCCACAGACCCCGTGTGC	1680
Db	1621		
Qy	1681	GCCTGGAGCTTCGGCTTGATGCTGTGTGGCCACCGCGGCGAGCACCGCGGATGGTT	1740
Db	1681		
Qy	1681	GCCTGGAGCTTCGGCTTGATGCTGTGTGGCCACCGCGGCGAGCACCGCGGATGGTT	1740
Db	1681		
Qy	1741	CAAGATATAGAGTCAGCGATGCTCTCTTTGTGCCAAAGAACTGGAGAAATCCC	1800
Db	1741		
Qy	1741	CAAGATATAGAGTCAGCGATGCTCTCTTTGTGCCAAAGAACTGGAGAAATCCC	1800
Db	1741		
Qy	1801	GTAGTGTGAAGTTCGGTGGCTACTGTGGGCCACGTGGTCTCTGCCATGTTCCCCAGT	1860
Db	1801		
Qy	1801	GTAGTGTGAAGTTCGGTGGCTACTGTGGGCCACGTGGTCTCTGCCATGTTCCCCAGT	1860
Db	1801		
Qy	1861	TCGTCTGGGCATCTCTGTGTGGCACACCGCCAGTGGAGTGACTCGCTCACTCCCCGG	1920
Db	1861		
Qy	1861	TCGTCTGGGCATCTCTGTGTGGCACACCGCCAGTGGAGTGACTCGCTCACTCCCCGG	1920
Db	1861		
Qy	1921	AGGGATGGACTAGAGTGTGTGTGACCCAGGGGCCATGGGGCTTATGCTTCCGAGTGT	1980
Db	1921		
Qy	1921	AGGGATGGACTAGAGTGTGTGTGACCCAGGGGCCATGGGGCTTATGCTTCCGAGTGT	1980
Db	1921		
Qy	1981	CAGAGGGTGGAGCCCGCCGCTGTGTGCTTATAGCTTGGTGTGGGCGACCCAGCGG	2040
Db	1981		
Qy	1981	CAGAGGGTGGAGCCCGCCGCTGTGTGCTTATAGCTTGGTGTGGGCGACCCAGCGG	2040
Db	1981		
Qy	2041	GGACCTCAAAACCGGGCCACACCGTGTGGGGCTGGATTGGTTGGCTTCTCTCGGT	2100
Db	2041		
Qy	2041	GGACCTCAAAACCGGGCCACACCGTGTGGGGCTGGATTGGTTGGCTTCTCTCGGT	2100
Db	2041		

Qy	2101	GTCTTCGACGATCCCTCACTCTCTCCCTGATTTGGTGC	CGCGTCA	GAGCGT	CGGCACAG	2160
Db	2101	GTCTTCGACGATCCCTCACTCTCCCTGATTTGGTGC	CGCGTCA	GAGCGT	CGGCACAG	2160
Qy	2161	AGGAGGCTTCTAGCTAGAGACAAGGTGGGCTTAGATCT	TGGGGCT	CCACCTT	CTGGGACC	2220
Db	2161	AGGAGGCTTCTAGCTAGAGACAAGGTGGGCTTAGATCT	TGGGGCT	CCACCTT	CTGGGACC	2220
Qy	2221	ACAAGCTATATAGTCAGGACCCCTCCCTCTCCCTTGC	CTTCCCT	GAAGAT	GAACGGCT	2280
Db	2221	ACAAGCTATATAGTCAGGACCCCTCCCTCTCCCTTGC	CTTCCCT	GAAGAT	GAACGGCT	2280
Qy	2281	CTGGGTAAAGCGGGGACGTGGTTTGGTGGCTTCCCT	TCCACCTT	CTCTGCT	TGGATCTTGC	2340
Db	2281	CTGGGTAAAGCGGGGACGTGGTTTGGTGGCTTCCCT	TCCACCTT	CTCTGCT	TGGATCTTGC	2340
Qy	2341	CCAAGCCACGCGACATCCGGCTCACTTGGGGCGCCT	TAGCCAC	GTGTGAT	GAGACTCC	2400
Db	2341	CCAAGCCACGCGACATCCGGCTCACTTGGGGCGCCT	TAGCCAC	GTGTGAT	GAGACTCC	2400
Qy	2401	ATCTAAAGCGGGGAAAATGACTTCCGAGCCATGAG	CAGTCTCT	TGGAAC	TAGTGGCTTACC	2460
Db	2401	ATCTAAAGCGGGGAAAATGACTTCCGAGCCATGAG	CAGTCTCT	TGGAAC	TAGTGGCTTACC	2460
Qy	2461	AGACCATGATCATGGCTGCTCCTTTCTCTTGGAGT	CTGTGCT	TTCACAT	TAGTGTCTG	2520
Db	2461	AGACCATGATCATGGCTGCTCCTTTCTCTTGGAGT	CTGTGCT	TTCACAT	TAGTGTCTG	2520
Qy	2521	TCCTCTGGACCTGGACCTGGCCCTTGGCCAGATT	CTCTGATT	CTCAT	GAGAGATCAACCT	2580
Db	2521	TCCTCTGGACCTGGACCTGGCCCTTGGCCAGATT	CTCTGATT	CTCAT	GAGAGATCAACCT	2580
Qy	2581	GTAACCTTCTGCGATGACCTTGTCTTTGGGCCAT	CATCAGCTT	TGGGGT	GGAGTAGAGAC	2640
Db	2581	GTAACCTTCTGCGATGACCTTGTCTTTGGGCCAT	CATCAGCTT	TGGGGT	GGAGTAGAGAC	2640
Qy	2641	ATAGGCCCCGGAAGGAATCAGTGTGGAGGTAGT	TGGGGCGT	TGTGTC	CGCTCGCTTCT	2700
Db	2641	ATAGGCCCCGGAAGGAATCAGTGTGGAGGTAGT	TGGGGCGT	TGTGTC	CGCTCGCTTCT	2700
Qy	2701	GTGGTGCTGTATGATTTCACGCTCTGCTGAC	TCTGCGG	AGCCAT	GATCCCCGCTGACTGC	2760
Db	2701	GTGGTGCTGTATGATTTCACGCTCTGCTGAC	TCTGCGG	AGCCAT	GATCCCCGCTGACTGC	2760
Qy	2761	CTTGAGATCTCTCCCACTCAGTTTCCCCCTGCT	TGCTGCG	AAGAGT	GTGTCTATACACTG	2820
Db	2761	CTTGAGATCTCTCCCACTCAGTTTCCCCCTGCT	TGCTGCG	AAGAGT	GTGTCTATACACTG	2820
Qy	2821	GTGTGCTTGAAGGCTTGTCATGTGTGCATG	GACGAC	AGGGCCG	TGCTCGGTGCTT	2880
Db	2821	GTGTGCTTGAAGGCTTGTCATGTGTGCATG	GACGAC	AGGGCCG	TGCTCGGTGCTT	2880
Qy	2881	TGGGGAGTCGGAGAAAGGTTTGGATTGGGGG	ACAACCTT	AACCGT	CGGTAGCCAGTGAGG	2940
Db	2881	TGGGGAGTCGGAGAAAGGTTTGGATTGGGGG	ACAACCTT	AACCGT	CGGTAGCCAGTGAGG	2940
Qy	2941	GAACACATATGCCCCCTCCCATCACCCAC	CAGCGCTCTT	TAACTT	TGACAAAGTCCC	3000
Db	2941	GAACACATATGCCCCCTCCCATCACCCAC	CAGCGCTCTT	TAACTT	TGACAAAGTCCC	3000
Qy	3001	AAAGTGACCTTCTGGGTGGGAAGGGAGC	AGGACATG	TGGCCCC	CGCTTCCCTGTGTC	3060
Db	3001	AAAGTGACCTTCTGGGTGGGAAGGGAGC	AGGACATG	TGGCCCC	CGCTTCCCTGTGTC	3060
Qy	3061	TTTCCCCTTCTGGCTGCCAACCTCTGCGT	CGCCACCG	CTGGCTT	CCCCGGCTGGAGTGG	3120
Db	3061	TTTCCCCTTCTGGCTGCCAACCTCTGCGT	CGCCACCG	CTGGCTT	CCCCGGCTGGAGTGG	3120
Qy	3121	AGGCTAGTCTCTGTCTTGGTTTCCATTTAAAT	TGAACCTT	CACAAC	ATCTCTAAATATT	3180
Db	3121	AGGCTAGTCTCTGTCTTGGTTTCCATTTAAAT	TGAACCTT	CACAAC	ATCTCTAAATATT	3180
Qy	3181	GGGGGATGACAAATGACTTTTTTCCCC	CAAAAAGT	GTGTAG	GAAATACAGCAGGTTAA	3240

[illegible]

RESULT 2

```

US/09-284-180A-1
; Sequence 1, Application US/09284180A
; GENERAL INFORMATION:
; APPLICANT: KIMURA, Toru
; APPLICANT: KIKUCHI, Kaoru
; TITLE OF INVENTION: NOVEL SEMAPHORIN W
; FILE REFERENCE: 0020-4546P
; CURRENT APPLICATION NUMBER: US/09/284,180A
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4008
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:

```

	; NAME/KEY: 5'UTR	
	; LOCATION: (1)..(75)	
	; FEATURE:	
	; NAME/KEY: misc_feature	
	; LOCATION: (76)..(2406)	
	; OTHER INFORMATION: Coding region	
	; FEATURE:	
	; NAME/KEY: 3'UTR	
	; LOCATION: (2407)..(3977)	
	; FEATURE:	
	; NAME/KEY: polyA_signal	
	; LOCATION: (3978)..(4008)	
	; FEATURE:	
	; NAME/KEY: misc_feature	
	; LOCATION: (1)..(4008)	
	; OTHER INFORMATION: strandedness: double	
	us-09-284-180A-1	
	Query Match 100.0%; Score 4008; DB 18; Length 4008;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 4008; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 GCCAGAGCCCGCGCAGTAGGGTACTAACTAGAGGCTGCTGGAGGCCGCCACCCCGGCAC	60
Db	1 GCCAGAGCCCGCGCAGTAGGGTACTAACTAGAGGCTGCTGGAGGCCGCCACCCCGGCAC	60
Qy	61 CAGCGCAGGCAGAGATGCTTGCCAGGCGCGAGCGGCCGCCGCCGCCGCCGCCGCCCT	120
Db	61 CAGCGCAGGCAGAGATGCTTGCCAGGCGCGAGCGGCCGCCGCCGCCGCCGCCGCCCT	120
Qy	121 CCGGTCTTTCCCTTCGCCGCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	180
Db	121 CCGGTCTTTCCCTTCGCCGCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	180
Qy	181 CCGGTGTCGGGCGCGCTGCCCGCTCAGTGCCCGCAGAACCTCGCTGCCCATCTCCGAGGCT	240
Db	181 CCGGTGTCGGGCGCGCTGCCCGCTCAGTGCCCGCAGAACCTCGCTGCCCATCTCCGAGGCT	240
Qy	241 GACTCCTATCTCACCGGTTTGACGGTCTCATACGTACAAATTACTGTCTCTCCTTTGTG	300
Db	241 GACTCCTATCTCACCGGTTTGACGGTCTCATACGTACAAATTACTGTCTCTCCTTTGTG	300
Qy	301 GATCCTGCCTCCCACACACTTTACGTGCGTGCACGGGATAGCATCTTCGCTTTAACCCCTC	360
Db	301 GATCCTGCCTCCCACACACTTTACGTGCGTGCACGGGATAGCATCTTCGCTTTAACCCCTC	360
Qy	361 CCTTCTCTGGGAAGAACCCCGAAGATCGACTGATGTACTCGTAGACTCACACAG	420
Db	361 CCTTCTCTGGGAAGAACCCCGAAGATCGACTGATGTACTCGTAGACTCACACAG	420
Qy	421 AACTGCAGGAGAAGGCAAGAGAGGAGCAATGTCACAATTTTATCCAGATTCCTCGCC	480
Db	421 AACTGCAGGAGAAGGCAAGAGAGGAGCAATGTCACAATTTTATCCAGATTCCTCGCC	480
Qy	481 ATTGTCAATGCCCTCTCACCTCCTCAGTGTGGGCACCCTTCGCTTTTGTATCCGAAGTCGGG	540
Db	481 ATTGTCAATGCCCTCTCACCTCCTCAGTGTGGGCACCCTTCGCTTTTGTATCCGAAGTCGGG	540
Qy	541 GTTAATTGATGTGCCAGTTTTCACGACAGTTTGAAGACTTGAAGACGCGCGGGGGAATGT	600
Db	541 GTTAATTGATGTGCCAGTTTTCACGACAGTTTGAAGACTTGAAGACGCGCGGGGGAATGT	600
Qy	601 CCTTTTCAGCCAGCTCAACGGTCAGCAGCTGTAAATGGCTGGGGGCGCTCTCTACACCGCC	660
Db	601 CCTTTTCAGCCAGCTCAACGGTCAGCAGCTGTAAATGGCTGGGGGCGCTCTCTACACCGCC	660
Qy	661 ACTGTGAAGAACTTCCTGGGACTTGAGCCCATCATCTCCGAGCTGTGGGTCCAGCTGAG	720
Db	661 ACTGTGAAGAACTTCCTGGGACTTGAGCCCATCATCTCCGAGCTGTGGGTCCAGCTGAG	720
Qy	721 GACTGGATTGGAACAGAGACCTTGTCACTCCTGCTTAAATGCTCCAGCCTTTGTGCGAGCT	780
Db	721 GACTGGATTGGAACAGAGACCTTGTCACTCCTGCTTAAATGCTCCAGCCTTTGTGCGAGCT	780

Qy	781	ATGTCCTGAGCCAGCTGAGTGGGGGATGAAGATGAGACGATGAATCTTTTTC	840	Qy	1861	TCCTCCCTGGGCATCTGTGTGTGACACAGCCAGTGGAGTGGCTGCTCCTCCCGG	1920
Db				Db			
Qy	781	ATGTCCTGAGCCAGCTGAGTGGGGGATGAAGATGAGACGATGAATCTTTTTC	840	Qy	1861	TCCTCCCTGGGCATCTGTGTGTGACACAGCCAGTGGAGTGGCTGCTCCTCCCGG	1920
Db				Db			
Qy	841	TTACGAGAGACTCCCGAGTGTGGACTCTATGAGCGCATCAAGTGCACAGAGTGGCC	900	Qy	1921	AGGATGGACTAGAGTGGTGGTACCCAGGGGCCATGGGGCTTATGCTTGGAGTGT	1980
Db				Db			
Qy	841	TTACGAGAGACTCCCGAGTGTGGACTCTATGAGCGCATCAAGTGCACAGAGTGGCC	900	Qy	1921	AGGATGGACTAGAGTGGTGGTACCCAGGGGCCATGGGGCTTATGCTTGGAGTGT	1980
Db				Db			
Qy	901	CGAGTGTGGGGGACCTTGGGGGAGGAGACCTTACAGAGAGATGACACAGCTTT	960	Qy	1981	CAGGAGGTGGAGCGCCGCTGCTTATAGCTTGTGTGGGGCAGCAGCGG	2040
Db				Db			
Qy	901	CGAGTGTGGGGGACCTTGGGGGAGGAGACCTTACAGAGAGATGACACAGCTTT	960	Qy	1981	CAGGAGGTGGAGCGCCGCTGCTTATAGCTTGTGTGGGGCAGCAGCGG	2040
Db				Db			
Qy	961	CTGAAGCTGACCTGTGTGTCAGGCGCCGAGCATGCGGGCTTCTGCGAG	1020	Qy	2041	GGACCTCAAAACCGGGGCCACACCGCTTGTGGGGCTGGATTGGCTTCTCTGGGT	2100
Db				Db			
Qy	961	CTGAAGCTGACCTGTGTGTCAGGCGCCGAGCATGCGGGCTTCTGCGAG	1020	Qy	2041	GGACCTCAAAACCGGGGCCACACCGCTTGTGGGGCTGGATTGGCTTCTCTGGGT	2100
Db				Db			
Qy	1021	GCTATGGCAGAGCTTGGGCTCAGCCTGGAGCGGAGACCCCATCTTTATGGATCTTT	1080	Qy	2101	GTTCTTGACAGATCCCTCAGCTCTCTCTGATTTGGTGGCGCTCAGCAGCGTGGGACAG	2160
Db				Db			
Qy	1021	GCTATGGCAGAGCTTGGGCTCAGCCTGGAGCGGAGACCCCATCTTTATGGATCTTT	1080	Qy	2101	GTTCTTGACAGATCCCTCAGCTCTCTCTGATTTGGTGGCGCTCAGCAGCGTGGGACAG	2160
Db				Db			
Qy	1081	TCCCTCCAGTGGAGAGCTGCCATCTCTGTGTGTGCTTCCGACCCCAAGACATC	1140	Qy	2161	AGGAGCTTCTAGCTAGACAAAGTGGCTTAGATCTGGGGCTTCCACCTTCTGGGACC	2220
Db				Db			
Qy	1081	TCCCTCCAGTGGAGAGCTGCCATCTCTGTGTGTGCTTCCGACCCCAAGACATC	1140	Qy	2161	AGGAGCTTCTAGCTAGACAAAGTGGCTTAGATCTGGGGCTTCCACCTTCTGGGACC	2220
Db				Db			
Qy	1141	CGGCGAGTGTGAATGTGCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCCT	1200	Qy	2221	ACAAAGCTATAGTACGAGCCCTCCCTCTCTGCGCTGAAGATGAACGGCTGCCCTGGCC	2280
Db				Db			
Qy	1141	CGGCGAGTGTGAATGTGCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCCT	1200	Qy	2221	ACAAAGCTATAGTACGAGCCCTCCCTCTCTGCGCTGAAGATGAACGGCTGCCCTGGCC	2280
Db				Db			
Qy	1201	GTCATGACACAGAGTGGCCCGCCAGCCGAGAGTGCATGCGCAACATGAAG	1260	Qy	2281	CTGGGTAAAGCGGGCAGTGGTGGTGGCTTCCCTCCACCTTCTCTGGATCTTGGC	2340
Db				Db			
Qy	1201	GTCATGACACAGAGTGGCCCGCCAGCCGAGAGTGCATGCGCAACATGAAG	1260	Qy	2281	CTGGGTAAAGCGGGCAGTGGTGGTGGCTTCCCTCCACCTTCTCTGGATCTTGGC	2340
Db				Db			
Qy	1261	CTCCAGAGTTGGATCTCACTCTCCCTGCGACACCGCTGCTCAGAGAC	1320	Qy	2341	CCAAAGCCAGCCACATCCGGCTCAGTGGGGCTCTTAGCCAGTGTGATGAGACCTCC	2400
Db				Db			
Qy	1261	CTCCAGAGTTGGATCTCACTCTCCCTGCGACACCGCTGCTCAGAGAC	1320	Qy	2341	CCAAAGCCAGCCACATCCGGCTCAGTGGGGCTCTTAGCCAGTGTGATGAGACCTCC	2400
Db				Db			
Qy	1321	CACCTCTCATGAGACGCGCTGTCCCGCTGACGGCCGCCCTGCTGGTCACTACA	1380	Qy	2401	ATCTAAAGCCGGGAAATGACTGCGACCATGAGCAGTCTCTGGAATAGTGGGTACCA	2460
Db				Db			
Qy	1321	CACCTCTCATGAGACGCGCTGTCCCGCTGACGGCCGCCCTGCTGGTCACTACA	1380	Qy	2401	ATCTAAAGCCGGGAAATGACTGCGACCATGAGCAGTCTCTGGAATAGTGGGTACCA	2460
Db				Db			
Qy	1381	GATACAGCTATCTCAGAGTGGTGGCCAGAGGTGACACGCTCTCAGGAAAGATAT	1440	Qy	2461	AGACCATGATCATGGCTGCTCTCTCTGGAGTCTGTGTTCACACATTAGTGTCTG	2520
Db				Db			
Qy	1381	GATACAGCTATCTCAGAGTGGTGGCCAGAGGTGACACGCTCTCAGGAAAGATAT	1440	Qy	2461	AGACCATGATCATGGCTGCTCTCTCTGGAGTCTGTGTTCACACATTAGTGTCTG	2520
Db				Db			
Qy	1441	GACGTGCTTACCTGGGACAGAGATGGACACCTCCACCGGCTGTCGCAATTGGAGCT	1500	Qy	2521	TCCTCTGACCTGGACCTGGCTTGGCCAGATCTCTGATCTCATGAGATCAACCT	2580
Db				Db			
Qy	1441	GACGTGCTTACCTGGGACAGAGATGGACACCTCCACCGGCTGTCGCAATTGGAGCT	1500	Qy	2521	TCCTCTGACCTGGACCTGGCTTGGCCAGATCTCTGATCTCATGAGATCAACCT	2580
Db				Db			
Qy	1501	CAGTCTAGTGTGGAGGATCTGGCCTTGTCCAGAACACAGCCGGTTGAGAGCATG	1560	Qy	2581	GTAACTTCTCGGATGGCTCTTGTCTTGGGCCATCAGCTTGGGGTGGAGTAAGGAC	2640
Db				Db			
Qy	1501	CAGTCTAGTGTGGAGGATCTGGCCTTGTCCAGAACACAGCCGGTTGAGAGCATG	1560	Qy	2581	GTAACTTCTCGGATGGCTCTTGTCTTGGGCCATCAGCTTGGGGTGGAGTAAGGAC	2640
Db				Db			
Qy	1561	AAATGTACACAGATGGCTGCTGGTGGCTCCATGAGGTGACACAGTGAACACC	1620	Qy	2641	ATAGCCCGGAAAGGGAATCAGTGTGGAGTGTGGGGCTGTGTGCCCTGGCTCTT	2700
Db				Db			
Qy	1561	AAATGTACACAGATGGCTGCTGGTGGCTCCATGAGGTGACACAGTGAACACC	1620	Qy	2641	ATAGCCCGGAAAGGGAATCAGTGTGGAGTGTGGGGCTGTGTGCCCTGGCTCTT	2700
Db				Db			
Qy	1621	AGCAACTGTGGCGCTTCCAGAGTGTCTGAGGTGTATCTTGGCCAGGACCCGCTGTC	1680	Qy	2701	GTGTGGCTGTATGATTTCCAGCTGTCTGACTCTTGGGGAGCGCATGATCCCTGACTGC	2760
Db				Db			
Qy	1621	AGCAACTGTGGCGCTTCCAGAGTGTCTGAGGTGTATCTTGGCCAGGACCCGCTGTC	1680	Qy	2701	GTGTGGCTGTATGATTTCCAGCTGTCTGACTCTTGGGGAGCGCATGATCCCTGACTGC	2760
Db				Db			
Qy	1681	GCCTGGAGCTTCCGCTTGTGTGTGTGGCCAGCGCGGAGCAGCGGGATGGTT	1740	Qy	2761	CTTGAGATCTCTCCCACTAGTTTCCCTTGTCTTGGAGAGTGTGTCTATACATG	2820
Db				Db			
Qy	1681	GCCTGGAGCTTCCGCTTGTGTGTGTGGCCAGCGCGGAGCAGCGGGATGGTT	1740	Qy	2761	CTTGAGATCTCTCCCACTAGTTTCCCTTGTCTTGGAGAGTGTGTCTATACATG	2820
Db				Db			
Qy	1741	CAAGATATAGAGTACAGCGGATGTCTCTTTGTGTCCAAAAGAACCTGGAGACATCCC	1800	Qy	2821	GTGTGCCCTAGAGGCTTCCATGTGTGCATGGAGACAGGGCCGGTCCCTGGGTGCTTT	2880
Db				Db			
Qy	1741	CAAGATATAGAGTACAGCGGATGTCTCTTTGTGTCCAAAAGAACCTGGAGACATCCC	1800	Qy	2821	GTGTGCCCTAGAGGCTTCCATGTGTGCATGGAGACAGGGCCGGTCCCTGGGTGCTTT	2880
Db				Db			
Qy	1801	GTAGTGTGAGTTCCGGTGGCTTACGTGGGCGACGTCCTGCGCATGTTCCCGCAGT	1860	Qy	2881	TGGGAGTCGCGAGAGAAAGTTGGAATGGGGGAGAACTTAACCTTCGGTAGCAGTGAAG	2940
Db				Db			
Qy	1801	GTAGTGTGAGTTCCGGTGGCTTACGTGGGCGACGTCCTGCGCATGTTCCCGCAGT	1860	Qy	2881	TGGGAGTCGCGAGAGAAAGTTGGAATGGGGGAGAACTTAACCTTCGGTAGCAGTGAAG	2940
Db				Db			
Qy	1861	GTAGTGTGAGTTCCGGTGGCTTACGTGGGCGACGTCCTGCGCATGTTCCCGCAGT	1920	Qy	2941	GTAGTGTGAGTTCCCGTGGCTTACATCAGCCACAGCGCTTCTTTAACTTTGAGCAAGTTCCC	3000
Db				Db			

Db	721	GCTGAGTGGGGGATGAAGATGGAGACGATGAATCTTTTTCCTTTCACCGGAGACCTCC	780
Qy	856	CGAGTGTGGACTCCTATGACGCGCATCAAGGTCCCAAGATGGGCCAGTGTGTGCGGGG	915
Db	781	CGAGTGTGGACTCCTATGAGGCGATCAAGGTCCCAAGATGGGCCAGTGTGTGCGGGG	840
Qy	916	GACCTTGGGGCAGGAAGACCTTCACGAGAGATGGACGAGTTCCTGAAGGCTGACCTG	975
Db	841	GACCTTGGGGCAGGAAGACCTTCACGAGAGATGGACGAGTTCCTGAAGGCTGACCTG	900
Qy	976	CTGTGCCACAGGCCCGAGCAGTGC CGGGCTCCGGGTTCGCAAGGCTATGGCAGAGCTT	1035
Db	901	CTGTGCCACAGGCCCGAGCAGTGC CGGGCTCCGGGTTCGCAAGGCTATGGCAGAGCTT	960
Qy	1036	CGGCTCAGCCTGGAGCGGGAACCCCATCTTTTATGGGATCTTTTCTCCCACTGGGAA	1095
Db	961	CGGCTCAGCCTGGAGCGGGAACCCCATCTTTTATGGGATCTTTTCTCCCACTGGGAA	1020
Qy	1096	GGAGCTGCCATCTCTGTGTGTGTGTCCTTCGGACCCCAAGACATCCGGGCAGTGTGAAT	1155
Db	1021	GGAGCTGCCATCTCTGTGTGTGTGTCCTTCGGACCCCAAGACATCCGGGCAGTGTGAAT	1080
Qy	1156	GGTCCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCTGTATGGACAACGAG	1215
Db	1081	GGTCCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCTGTATGGACAACGAG	1140
Qy	1216	GTGCCCCAGCCAGACCTGGAGAGTGCATCGCAACAACTGAAGCTCCAGCAGTTTGGGA	1275
Db	1141	GTGCCCCAGCCAGACCTGGAGAGTGCATCGCCNAACTGAAGCTCCAGCAGTTTGGGA	1200
Qy	1276	TCCTCACTCTCCCTGCCAGACCGGTGCTCACCTTTATCAGACACCACTCTCATGTGAC	1335
Db	1201	TCCTCACTCTCCCTGCCAGACCGGTGCTCACCTTTATCAGACACCACTCTCATGTGAC	1260
Qy	1336	AGGCCGTGTTCGGGCTGACGGCCGCCCTGCTGCTCACTACAGATACAGCCTATCTC	1395
Db	1261	AGGCCGTGTTCGGGCTGACGGCCGCCCTGCTGCTCACTACAGATACAGCCTATCTC	1320
Qy	1396	AGAGTCTGTGCCACAGGGTGACCAGCCTCTCAGGGAAAGAAATATGAGTGTCTACCTG	1455
Db	1321	AGAGTCTGTGCCACAGGGTGACCAGCCTCTCAGGGAAAGAAATATGAGTGTCTACCTG	1380
Qy	1456	GGACAGAGGATGGACACCTCCACCGGGCTGTGCGCATTTGGAGCTCAGCTCAGTGTCTG	1515
Db	1381	GGACAGAGGATGGACACCTCCACCGGGCTGTGCGCATTTGGAGCTCAGCTCAGTGTCTG	1440
Qy	1516	GAGGATCTGGCTTGTTCACAGAACACACCGCGTTGAGAGCATGAATTTGTACCAGCAT	1575
Db	1441	GAGGATCTGGCTTGTTCACAGAACACACCGCGTTGAGAGCATGAATTTGTACCAGCAT	1500
Qy	1576	TGGCTCTGTGGGCTCCCATCTAGAGTGACACACAGTGAACACCACTGTGGCCGT	1635
Db	1501	TGGCTCTGTGGGCTCCCATCTAGAGTGACACACAGTGAACACCACTGTGGCCGT	1560
Qy	1636	CTCAGAGCTGCTCGGAGTGTATCTTGCCCGCAGACCCCGTGTGCGCTGGAGCTTCCGG	1695
Db	1561	CTCAGAGCTGCTCGGAGTGTATCTTGCCCGCAGACCCCGTGTGCGCTGGAGCTTCCGG	1620
Qy	1696	CTTGATGCTTGTGTGGCCACGCGGGCGAGCACCGGGGATGGTTCAAGATATAGAGTCA	1755
Db	1621	CTTGATGCTTGTGTGGCCACGCGGGCGAGCACCGGGGATGGTTCAAGATATAGAGTCA	1680
Qy	1756	GCGATCTCTCTTTTGTCTCAAAAGAACCTGGAGAACATCCGTAGTGTGTGAGTT	1815
Db	1681	GCGATGCTCTCTTTTGTCTCAAAAGAACCTGGAGAACATCCGTAGTGTGTGAGTT	1740
Qy	1816	CCGGTGGCTACTGTGGGCCAGTGGTCTGCCATGTTCCCGCAGTTCTGCTGGGCATCC	1875
Db	1741	CCGGTGGCTACTGTGGGCCAGTGGTCTGCCATGTTCCCGCAGTTCTGCTGGGCATCC	1800
Qy	1876	TGTGTGGCACAGCCCACTGGAGTGACTGCGCTCACTCCCGGAGGGATGACATAGAG	1935
Db	1801	TGTGTGGCACAGCCCACTGGAGTGACTGCGCTCACTCCCGGAGGGATGACATAGAG	1860

Qy	1936	GTGTTGGTGACCCACGAGGGCCATCGGGGCTTATGCTTGCAGTGTCTCAGAGGGTGGAGCC	1995
Db	1861	GTGTTGGTGACCCACGAGGGCCATCGGGGCTTATGCTTGCAGTGTCTCAGAGGGTGGAGCC	1920
Qy	1996	GCCCGCTGGTGCTGCTTATAGCTTGGTGTGGGACGCCAGCGGGACCCCTCAAAACGG	2055
Db	1921	GCCCGCTGGTGCTGCTTATAGCTTGGTGTGGGACGCCAGCGGGACCCCTCAAAACGG	1980
Qy	2056	GCCACACCCGTTGTGGGGCTGGATGTGTGGCTTTCTCCTGGGTCTTGTGCAGCATCC	2115
Db	1981	GCCACACCCGTTGTGGGGCTGGATGTGTGGCTTTCTCCTGGGTCTTGTGCAGCATCC	2040
Qy	2116	CTCACTCTCCCTCCTGATTGGTCGCCGTACGACGCTCGGCACAGAGGGAGCTTCTAGCT	2175
Db	2041	CTCACTCTCCCTCCTGATTGGTCGCCGTACGACGCTCGGCACAGAGGGAGCTTCTAGCT	2100
Qy	2176	AGACACAAGGTGGCTTAGATCTCGGGGCTCCACCTTCTGGGACCCACAAGCTATAGTCAG	2235
Db	2101	AGACACAAGGTGGCTTAGATCTCGGGGCTCCACCTTCTGGGACCCACAAGCTATAGTCAG	2160
Qy	2236	GACCTTCCTCTCCTTCGCCTGAAGATGAACGGCTGCCCTTGGCCCTGGGTAAAGCGGGC	2295
Db	2161	GACCTTCCTCTCCTTCGCCTGAAGATGAACGGCTGCCCTTGGCCCTGGGTAAAGCGGGC	2220
Qy	2296	AGTGGTTTTGGTGGCTTCCCCTCCACCCCTTCTCGCTGGATTTCTTGCCTCAAGCCACGCCAC	2355
Db	2221	AGTGGTTTTGGTGGCTTCCCCTCCACCCCTTCTCGCTGGATTTCTTGCCTCAAGCCACGCCAC	2280
Qy	2356	ATCCGGTCTACTGGGGGCCCTCTTAGCCACGTGTGATGAGACCTCCATCTAA	2406
Db	2281	ATCCGGTCTACTGGGGGCCCTCTTAGCCACGTGTGATGAGACCTCCATCTAA	2331

RESULT 4

```

US-09-284-180A-2
; Sequence 2, Application US/09284180A
; GENERAL INFORMATION:
; APPLICANT: KIMURA, Toru
; APPLICANT: KIKUCHI, Kaoru
; TITLE OF INVENTION: NOVEL SEMAPHORIN GENE: SEMAPHORIN W
; FILE REFERENCE: 0020-4548P
; CURRENT APPLICATION NUMBER: US/09/284,180A
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2331
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2331)
; OTHER INFORMATION: Coding region
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2331)
; OTHER INFORMATION: strandedness: double
; US-09-284-180A-2

```

	Query Match	58.2%	Score 2331;	DB 18;	Length 2331;
	Best Local Similarity	100.0%;	Presd. No. 4.9e-307;		
	Matches 2331;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	76	ATGCTTGCAGGGCCGAGCGCCCGCCGGGGCCCCCGCGCGCCTCCGGCTCTTCCCTTC	135		
Db	1	ATGCTTCCAGGGCCGAGCGCCCGCCGGGGCCCCCGCGCTCTTCCCTTC	60		
Qy	136	CCGCGCGCGCTGTGCTGCTGCTGCTGCGGATACTAAGCGCCCCCGGTGCGCGCCGC	195		
Db	61	CCGCGCGCGCTGTGCTGCTGCTGCTGCGGATACTAAGCGCCCCCGGTGCGCGCCGC	120		
Qy	196	GTCCCCCGGCTCAGTGCCCGAGAACCTCGCTGCCATCTCCGAGGCTGACTCTATCTCACC	255		

Db 121 GTCCCCGCTCAGTGGCCGAACCTCGTCCCATCTCCGAGGCTGACTCCTATCTCAC 180
QY 256 CGGTTTGCAGCGTCTCATAGTACAATTACTCTGCTCTCCTTGTGGATCTCTGCTCCAC 315
Db 181 CGGTTTGCAGCGTCTCATAGTACAATTACTCTGCTCTCCTTGTGGATCTCTGCTCCAC 240
QY 316 ACACTTTACGTCGGTGACGGGATAGATCTTCGCTTTAAACCTCCCTCTCTCTGGGAA 375
Db 241 ACACTTTACGTCGGTGACGGGATAGATCTTCGCTTTAAACCTCCCTCTCTCTGGGAA 300
QY 376 AGACCCGGAAGGATCGACTGGATGATGCTGACACTCACAGACAGAACTGCAGAGAAA 435
Db 301 AGACCCGGAAGGATCGACTGGATGATGCTGACACTCACAGACAGAACTGCAGAGAAA 360
QY 436 GGCAAGAAAGAGGACGAATGTCAAAATTTTATCCAGATTCTCGCCATTGTCAATGCCTCT 495
Db 361 GGCAAGAAAGAGGACGAATGTCAAAATTTTATCCAGATTCTCGCCATTGTCAATGCCTCT 420
QY 496 CACCTCTCAGTGGCGCACTTCGCTTTTATCCGAAGTGCAGGCTTATGATGTGTC 555
Db 421 CACCTCTCAGTGGCGCACTTCGCTTTTATCCGAAGTGCAGGCTTATGATGTGTC 480
QY 556 AGTTTCCAGCAGGTTGAAAGACTTGAGAGCGCGGGGAAATGTCTTTTGAAGCCAGCT 615
Db 481 AGTTTCCAGCAGGTTGAAAGACTTGAGAGCGCGGGGAAATGTCTTTTGAAGCCAGCT 540
QY 616 CAACGGTTCAGCAGCTTAATGGCTGGGGCGTCTCTACACGCCCACTGTGAAGAACTTC 675
Db 541 CAACGGTTCAGCAGCTTAATGGCTGGGGCGTCTCTACACGCCCACTGTGAAGAACTTC 600
QY 676 CTGGGACTGAGCCCATCATCTCCAGCTGGGGTGGAGCTGAGACTGGATTCGAACA 735
Db 601 CTGGGACTGAGCCCATCATCTCCAGCTGGGGTGGAGCTGAGACTGGATTCGAACA 660
QY 736 GAGACCTTGTATCTGGCTTAATGCTCCAGCTTTGTCGAGCTATGCTGAGCCCA 795
Db 661 GAGACCTTGTATCTGGCTTAATGCTCCAGCTTTGTCGAGCTATGCTGAGCCCA 720
QY 796 GCTGAGTGGGGGATGAAGATGGAGACGATGAATCTTTTTCACGGAGACCTCC 855
Db 721 GCTGAGTGGGGGATGAAGATGGAGACGATGAATCTTTTTCACGGAGACCTCC 780
QY 856 CGAGCTTGTGACTCCTATGAGCGCATCAAGTCCCAAGAGTGCCCGAGTGTGTGGGG 915
Db 781 CGAGTGTGTGACTCCTATGAGCGCATCAAGTCCCAAGAGTGCCCGAGTGTGTGGGG 840
QY 916 GACCTTGGGGGAGGAGACCTTTCAGCAGATGGACGAGTTTCTGAAGGCTGACCTG 975
Db 841 GACCTTGGGGGAGGAGACCTTTCAGCAGATGGACGAGTTTCTGAAGGCTGACCTG 900
QY 976 CTGTGCCCCAGGGCCGAGCATGGCCGGGCTCCGGGGTTCTCGAGGCTATGGCAGAGCTT 1035
Db 901 CTGTGCCCCAGGGCCGAGCATGGCCGGGCTCCGGGGTTCTCGAGGCTATGGCAGAGCTT 960
QY 1036 CGGCTCAGCCTGGAGCGGAACCCCATCTTTTATGGATCTTTTCCCTCCAGTGGGAA 1095
Db 961 CGGCTCAGCCTGGAGCGGAACCCCATCTTTTATGGATCTTTTCCCTCCAGTGGGAA 1020
QY 1096 GGAGCTGCCATCTCTGCTGTGTGCTTCCGACCCCAAGACATCCGGGCACTGCTGAAT 1155
Db 1021 GGAGCTGCCATCTCTGCTGTGTGCTTCCGACCCCAAGACATCCGGGCACTGCTGAAT 1080
QY 1156 GGTCCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCCTGTCTATGGACACGAG 1215
Db 1081 GGTCCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCCTGTCTATGGACACGAG 1140
QY 1216 GTGCCCCAGCCAGACCTGGAGAGTGCATCGCCCAACAACATGAAGCTCAGCAGTTTGA 1275
Db 1141 GTGCCCCAGCCAGACCTGGAGAGTGCATCGCCCAACAACATGAAGCTCAGCAGTTTGA 1200
QY 1276 TCCTCACTCTCCCTGCCAGACCGCTGTCACTTTATCAGAGACACCCCTCTCATGGAC 1335

Db 1201 TCCTCACTCTCCCTGCCAGACCGCGTCTCACCTTTATCAGAGACCACCCCTCTCATGGAC 1260
QY 1336 AGCCCGCTGTTCGCCGCTGACGCCGCCCTCTGCTACTACAGATACAGCCTATCTC 1395
Db 1261 AGCCCGCTGTTCGCCGCTGACGCCGCCCTCTGCTACTACAGATACAGCCTATCTC 1320
QY 1396 AGAGTCTGTGGCCACAGGGTGACAGCCTCTCAGGAAAGAAATATGACGTGCTCTACCTG 1455
Db 1321 AGAGTCTGTGGCCACAGGGTGACAGCCTCTCAGGAAAGAAATATGACGTGCTCTACCTG 1380
QY 1456 GGGACAGAGGATGAGACACCTCCACCGGGCTGCGGATTTGGAGCTCAGCTCAGTGTCTTG 1515
Db 1381 GGGACAGAGGATGAGACACCTCCACCGGGCTGCGGATTTGGAGCTCAGCTCAGTGTCTTG 1440
QY 1516 GAGGATCTGGCCTTGTTCAGAAACACAGCCGGTTGAGAGCATGAAATTTGTACCACGAT 1575
Db 1441 GAGGATCTGGCCTTGTTCAGAAACACAGCCGGTTGAGAGCATGAAATTTGTACCACGAT 1500
QY 1576 TGGCTCCTGTGGGCTCCCATACTGAGGTGACAAAGTGAACACCAAGCAACTGTGCGCT 1635
Db 1501 TGGCTCCTGTGGGCTCCCATACTGAGGTGACAAAGTGAACACCAAGCAACTGTGCGCT 1560
QY 1636 CTCCAGAGCTGCTCGGAGTGTATCTTGGCCAGGACCCCGTGTGCGCTGGAGCTTCCGG 1695
Db 1561 CTCCAGAGCTGCTCGGAGTGTATCTTGGCCAGGACCCCGTGTGCGCTGGAGCTTCCGG 1620
QY 1696 CTGTAGTCTGTGTGGCCACCGCGGAGACCGCGGATGTTCAAGATATAGAGTCA 1755
Db 1621 CTGTAGTCTGTGTGGCCACCGCGGAGACCGCGGATGTTCAAGATATAGAGTCA 1680
QY 1756 GGGGATGTCTCTTCTTTTGTCCAAAGAACCTTGGAGAACATCCCGTAGTGTTTGAAGTT 1815
Db 1681 GGGGATGTCTCTTCTTTGTCCAAAGAACCTTGGAGAACATCCCGTAGTGTTTGAAGTT 1740
QY 1816 CGGCTGCTACTGTGGCCAGCTGTGCTTCCCATGTTCCCGGAGTTCCTGCTGGGATCC 1875
Db 1741 CGGCTGCTACTGTGGCCAGCTGTGCTTCCCATGTTCCCGGAGTTCCTGCTGGGATCC 1800
QY 1876 TGTGTGTGCACAGACCCAGTGGAGTGTGCTTCCCTCAGTCCCGGAGGATGAGCTAGAG 1935
Db 1801 TGTGTGTGCACAGACCCAGTGGAGTGTGCTTCCCTCAGTCCCGGAGGATGAGCTAGAG 1860
QY 1936 GTGGTGGTGAACCCAGGGCCATGGGGCTTATGCTTGGAGTGTTCAGGAGGTGGAGCC 1995
Db 1861 GTGGTGGTGAACCCAGGGCCATGGGGCTTATGCTTGGAGTGTTCAGGAGGTGGAGCC 1920
QY 1996 GCGCGGTGCTGCTCTTATAGCTTGGTGTGGGCGACGCGGGGACCCCTCAAAACCGG 2055
Db 1921 GCGCGGTGCTGCTCTTATAGCTTGGTGTGGGCGACGCGGGGACCCCTCAAAACCGG 1980
QY 2056 GCCACACAGCTTGTGGGGCTGGATTGGCTTTCCTCTGGTGTCTTTCGAGCATCC 2115
Db 1981 GCCACACAGCTTGTGGGGCTGGATTGGCTTTCCTCTGGTGTCTTTCGAGCATCC 2040
QY 2116 CTCACCTCTCCTCTGATTTGGTTCGCGTACAGAGCTGTGGCGACAGAGGAGCTTTCAGT 2175
Db 2041 CTCACCTCTCCTCTGATTTGGTTCGCGTACAGAGCTGTGGCGACAGAGGAGCTTTCAGT 2100
QY 2176 AGAGACAGGTTGGGCTTAGATCTGGGGCTTCCACCTTTCGGGACCAAGCTATAGTCAG 2235
Db 2101 AGAGACAGGTTGGGCTTAGATCTGGGGCTTCCACCTTTCGGGACCAAGCTATAGTCAG 2160
QY 2236 GACCTCTCCCTCTCTCGCCTGAAGATGAACGGCTGCCCTGCCCTGGGTGAAGCGGGC 2295
Db 2161 GACCTCTCCCTCTCTCGCCTGAAGATGAACGGCTGCCCTGGGTGAAGCGGGC 2220
QY 2296 AGTGGTTTGTGGGCTTCCCTCCACCTTCTGCTGGATTTCTTGCCAAAGCCAGCCAC 2355
Db 2221 AGTGGTTTGTGGGCTTCCCTCCACCTTCTGCTGGATTTCTTGCCAAAGCCAGCCAC 2280
QY 2356 ATCCGGCTCACTGGGGCGCTCTAGCCAGCTGTGATGAGACCTCCATCTAA 2406
Db 2281 ATCCGGCTCACTGGGGCGCTCTAGCCAGCTGTGATGAGACCTCCATCTAA 2331

RESULT 5					
US-10-144-771-19015					
; Sequence 19015, Application US/10144771					
; GENERAL INFORMATION:					
; APPLICANT: VENTER, J. Craig					
; TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF					
; FILE REFERENCE: CL001321					
; CURRENT APPLICATION NUMBER: US/10/144,771					
; CURRENT FILING DATE: 2002-05-15					
; NUMBER OF SEQ ID NOS: 47235					
; SEQ ID NO 19015					
; LENGTH: 2334					
; TYPE: DNA					
; ORGANISM: HUMAN					
US-10-144-771-19015					
Query Match 53.4%; Score 2138.8; DB 46; Length 2334;					
Best Local Similarity 95.1%; Pred. No. 6.2e-281;					
Matches 2219; Conservative 0; Mismatches 112; Indels 3; Gaps 1					
QY	76	ATGCTTGCCAGGGCCGAGCGGCCGCCCGGGCCCCCGGGCGCCTCCGGTCTTTTCCTTC	135		
Db	1	ATGCTGCCAGGGCCGAGCGGCCGCCCGGGCCCCCGGGCGCCTCCGGTCTCTCTCTTC	60		
QY	136	CCGCGCGCGTGTGC--GCTGCTGCTGCTGCTGCGGATACTAAGGCGCCCGGTGTGCGGC	192		
Db	61	CCGCGCGCGTGTGCTGCTGCTGCTGCTGCGGATGCTGAGCGCCCGGTGTGCGGC	120		
QY	193	CGGTCGCCCGCTCAGTGCACGAACCTCGCTGCCCATCTCCGAGGCTGACTCCTATCTC	252		
Db	121	CGCGTCCCCCGCTCGGTGCCAGAACCTCGCTGCCCATCTCCGAGGCTGACTCCTATCTC	180		
QY	253	ACCGGGTTTCGAGGCTCATACGTACAATTACTCTGCTCCTCTTGCGGATCCTGCCCTCC	312		
Db	181	ACCGGGTTTCGGGCCCTCATACGTACAATTACTCTGCTCCTCTGCGGATCCTGCCCTCC	240		
QY	313	CACACACTTTACGTCGGTGCACGGGATAGACTCTTCGCTTTAAACCCTCCCCTTCTCTGGG	372		
Db	241	CACACACTTTACGTCGGTGCACGGGACAGCATCTTCGCTTTAAACCCTCCCCTTCTCTGGG	300		
QY	373	GAAAGACCCCGAAGGATCGACTGGATGGTAGCTACCTGAGACTCACAGACAGAAGTCGAGG	432		
Db	301	GAAAGCTCGAAGGATCGACTGGATGGTGCACGAGACTCACAGACAGAAGTCGAGGAG	360		
QY	433	AAAGGCAAGAAAGAGGACGAATGFCAAATTTATCCAGATTCGCGCATGTCTCAATGCC	492		
Db	361	AAAGGCNAGAAAGAGGACGAATGTCACAATTTATCCAGATTCGCGCATGTCCCAATGCC	420		
QY	493	TCTCAGCTCTCAGTGGCGACCTTCGCTTTTGGATCCGAAGTCGGGGTTATTGATGTG	552		
Db	421	TCTCAGCTCTCAGTGGCGACCTTCGCTTTTGGATCCGAAGTCGGGGTTATTGATGTG	480		
QY	553	TCCAGTTTCCAGCAGGTTGAAGACTTTGAGAGCGCGCGGGGAATGTCCTTTTCAGCCA	612		
Db	481	TCCAGTTTCCAGCAGGTTGAAGACTTTGAGAGCTTGAGAGTGGCGGGGAATGTCCTTTTCAGCCA	540		
QY	613	GCTCAACGGTCAAGCAGCTGTAATGGCTGGGGCGCTCCTCTACACCGGCACACTGTGAAGAAC	672		
Db	541	GCTCAACGGTCAAGCAGCTGTAATGGCTGGGGCGCTCCTCTACACCGGCACACTGTGAAGAAC	600		
QY	673	TTCTCTGGGACTGAGCCCATCATCTCCGAGCTGTGGGTGCAGCTGAGGACTGGATTGCA	732		
Db	601	TTCTCTGGGACAGACGGCGATTATCTCCGAGCTGTGGGTGCAGCTGAGGACTGGATTGCA	660		
QY	733	ACAGAGACCTTGTTCATCCTGGCTTAATGCTCCAGCCTTGTGCGAGCTATGGTCTGAGC	792		
Db	661	ACAGAGACCTTGTTCATCCTGGCTTAATGCCCCAGCCTTGTGCGAGCTATGGTCTGAGC	720		
QY	793	CCAGCTGAGTGGGGGATGAAGATGGAGACGATGAAATCTTTTTTCTTCACGGAGACC	852		
Db	721	CCGGCTCAGTGGGGGATGAAGATGGAGACGATGAAATCTTTTTTCTTCACGGAGACC	780		

QY 1933 GAGGTGGTGGTACCCAGGGCCATGGGGGCTTATGCTTGGAGTGTCCAGAGGGTGA 1992
DB 1861 GAGGTGGTGGTACCCAGGGCCATGGGGGCTTATGCTTGGAGTGTCCAGAGGGTGA 1920
QY 1993 GCGCCCGCGTGGTGGCTTATAGCTTGGTGGGCGAGCCAGCGGGGACCCCTCAAC 2052
DB 1921 GCGCCCGCGTGGTGGCTTATAGCTTGGTGGGCGAGCCAGAGGGGGCCCGCAAC 1980
QY 2053 CCGGGCCACACCGTGTGGGGCTGGATTGGTGGCTTCTCTGGGTGTTCTTCAGCA 2112
DB 1981 CCGGGCCACACCGTGTGGGGCTGGCTGGTGGCTTCTCTGGGTGTTCTACGCA 2040
QY 2113 TCCCTACACTCTCTCTGATGGTGGCTGAGCAGCGTGGCGAGAGGAGTCTTA 2172
DB 2041 TCTCTACACTCTCTCTGATGGTGGCTGAGCAGCGCGGCGAGAGGAGGCTTCTA 2100
QY 2173 GCTAGAGCAAGTGGGCTTAGATCTGGGGCTCCACCTTCTGGGACCACAGCTATAGT 2232
DB 2101 GCTAGAGCAAGTGGGCTTAGATCTGGGGCTCCACCTTCTGGGACCACAGCTATAGC 2160
QY 2233 CAGGACCCCTCCCTCTCCCTGAAGATGAACGGCTGCCCTGGCCCTGGGTAAAGCG 2292
DB 2161 CAAGACCCCTCCCTCTCCCTGAAGATGAACGGCTGCCCTGGCCCTGGGTAAAGCG 2220
QY 2293 GGCAGTGGTGGTGGCTTCCCTCCACCTTCTCTGGTGGATCTTGGCCAAAGCCAGCC 2352
DB 2221 GGCAGTGGTGGTGGCTTCCCTCCACCTTCTCTGGTGGATCTTGGCCAAAGCCAGCC 2280
QY 2353 CACATCCGGCTCACATGGGGCGCTCTAGCAGCTGTGATGAGACCTCCACTAA 2406
DB 2281 CACATCCGGCTCACATGGGGCTCCTCTGGCCACGTGGATGAGACGCTATCTAA 2334

RESULT 6
US-60-360-207-19015
; Sequence 19015, Application US/60360207
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
; FILE REFERENCE: CL001321
; CURRENT APPLICATION NUMBER: US/60/360, 207
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 47235
; SEQ ID NO 19015
; LENGTH: 2334
; TYPE: DNA
; ORGANISM: HUMAN
US-60-360-207-19015

Query Match 53.4%; Score 2138.8; DB 91; Length 2334;
Best Local Similarity 95.1%; Pred. No. 6.2e-281;
Matches 2219; Conservative 0; Mismatches 112; Indels 3; Gaps 1;

QY 76 ATGCTTGCAGGGCGAGGGCCCGCGGGCCCGCGGGCCCGCTCCGGTCTTCCCTTC 135
DB 1 ATGCTTGCAGGGCGAGGGCCCGCGGGCCCGCGGGCCCGCTCCGGTCTTCCCTTC 60
QY 136 CCGCCCGCGCTGTC---GCTGCTGCTGCTGCTGGGGTACTAAGCGCCCGGTGTGCGCG 192
DB 61 CCGCCCGCGCTGCTGCTGCTGCTGCTGCTGGGGTACTAAGCGCCCGGTGTGCGCG 120
QY 193 CCGGTCGCCCGCTCAGTGGCCAGAACCTCGTGGCCCATCTCCGAGGCTGACTTCATCTC 252
DB 121 CCGGTCGCCCGCTCGGTGCCAGAACCTCGTGGCCCATCTCCGAGGCTGACTTCATCTC 180
QY 253 ACCCGGTTTGCAGGCTCATACGATACAAATTAATCTGCTCTCTCTGATCCTGCCCTCC 312
DB 181 ACCCGGTTTGGCGGCCCTCATACGATACAAATTAATCTGCTCTCTCTGATCCTGCCCTCC 240
QY 313 CACACACTTACGTCGGTGCAGGGATAGCACTTCGGCTTTAAACCTGCCCTTCTCTGGG 372
DB 241 CACACACTTACGTCGGTGCAGGGATAGCACTTCGGCTTTAAACCTGCCCTTCTCTGGG 300

QY 373 GAAAGACCCCGAAGGATCGACTGGATGGTACTCGACTGACACTCACAGACAGAACTGACGGAAG 432
DB 301 GAAAGACCCCGAAGGATCGACTGGATGGTACTCGACTGACACTCACAGACAGAACTGACGGAAG 360
QY 433 AAAGCAAGAAAGAGGACGAATGTCACAAATTTATCCAGATTTCTGCCAATTTCTCAATGCC 492
DB 361 AAAGCAAGAAAGAGGACGAATGTCACAAATTTATCCAGATTTCTGCCAATTTCCCAATGCC 420
QY 493 TCTCACTCTCTCACGTGGGCGACCTTGCCTTTTGTATCCGAAGTGGGGGTTATTTGATGTG 552
DB 421 TCTCACTCTCTCACGTGGGCGACCTTGCCTTTTGTATCCGAAGTGGGGGTTATTTGATGTG 480
QY 553 TCCAGTTTCCACAGGTTGAAAGACTTGAGACGGCGCGGGGAAATGTCCTTTTTCAGCCA 612
DB 481 TCCAGTTTCCACAGGTTGAAAGACTTGAGACGGCGCGGGGAAATGTCCTTTTTCAGCCA 540
QY 613 GCTCAACGGTACGACGCTGTAATGGCTGGGGGCGCTCTCTACACGCCCACTGTGAAGAAC 572
DB 541 GCTCAACGGTACGACGCTGTAATGGCTGGGGGCGCTCTCTACACGCCCACTGTGAAGAAC 600
QY 673 TTCTCTGGGACTGAGCCCATCATCTCCGAGCTGTGGTTCGAGCTGAGGACTGGATTCCA 732
DB 601 TTCTCTGGGACAGACCGGATTAATCTCCGAGCTGTGGTTCGAGCTGAGGACTGGATTCCA 660
QY 733 ACAGAGACCTGTCTATCTCTGGCTTAATGCTCCAGCCTTTGTCGCACTATGTCCTGAGC 792
DB 661 ACAGAGACCTGTCTATCTCTGGCTTAATGCTCCAGCCTTTGTCGCACTATGTCCTGAGC 720
QY 793 CCAGCTGAGTGGGGGATGAAGATGGAGACGATGAATCTTTTTCCTTTCACGGAGACC 852
DB 721 CCAGCTGAGTGGGGGATGAAGATGGAGACGATGAATCTTTCTTTCCTTTCACGGAGACC 780
QY 853 TCCCGAGTGTGGACTCCTATGAGCGCATCAAGGTCCCAAGAGTGGCCGAGTGTGTGG 912
DB 781 TCCCGAGTGTGGACTCCTATGAGCGCATCAAGGTCCCAAGAGTGGCCGAGTGTGTGG 840
QY 913 GGGGACCTTTGGGGCAGGAAGACCCCTTACAGAGATGGACACGCTTTCTGAAGGCTGAC 972
DB 841 GGGGACCTTTGGGGCAGGAAGACCCCTTACAGAGATGGACACGCTTTCTGAAGGCTGAC 900
QY 973 CTGCTGTCCCGAGGCGCGAGCATGGCGGGCTCCGGGTTCTCGAGCTATGCGAGAG 1032
DB 901 CTGCTGTGTCCAGGCGCGAGCATGGAAGGCTCGGGGTTCTCGAGGATATGACAGAG 960
QY 1033 CTTTCGGCTTCAGCTTGGAGCGGAACCCCATCTTTTATGGATCTTTTCCTCCAGTGG 1092
DB 961 CTTTCAGCTTCAGCTTGGCGGGGACCCCTCTTTATGGCATCTTTTCCTCCAGTGG 1020
QY 1093 GAAGAGCTGCCATCTCTGCTGTGTGCTTCCGACCCCAAGAGATCCGGGCACTGTG 1152
DB 1021 GAAGAGCGCCCATTTCTGCTGTGTGCTTCCGACCCCAAGAGATCCGGGCACTGTG 1080
QY 1153 AATGCTCTCTTAGAGACTAAACATGACTGCAACAGGGGACTGCCTCTCATGACACAC 1212
DB 1081 AATGCTCTCTTAGAGACTAAACATGACTGCAACAGGGGACTGCCTCTCATGACACAC 1140
QY 1213 GAGGTGCCCGCCAGACCTGGAGAGTGCATTCGCAACAACTGAAGCTCCAGCAGTTT 1272
DB 1141 GAGGTGCCCGCCAGACCTGGAGAGTGCATTCACCAACACATGAGTTCCAGCAGTTT 1200
QY 1273 GGATCCTCACTCTCCTGCCAGACCGGCTGCTACCTTTATCAGAGACACCCCTCTCATG 1332
DB 1201 GGATCCTCACTCTCCTGCCAGACCGGCTGCTACCTTTATCAGAGACACCCCGCTCATG 1260
QY 1333 GACAGGCGGCTTCCCGGCTGACGCGCCCTGCTGGTCACTACAGATACAGCCTAT 1392
DB 1261 GACAGGCGGCTTTCACAGCTGATGGCGCCCTGCTGGTCACTACGAGATACAGCCTAT 1320
QY 1393 CTCAGAGCTGGCGCCACAGGCTGACAGCCTCTCAGGGAAGAAATATGACGTGCTTAC 1452
DB 1321 CTCAGAGTGTGGCTCACCGGCTGACAGCCTCTCAGGGAAGAAATATGATGTGCTTAC 1380
QY 1453 CTGGGACAGAGGATGGACACCTCCACCGGGCTGTGGCATTTGGAGCTCAGCTCAGTGTG 1512

```

Db 1381 CTGGGACAGAGGATGGGACCTCCTCATCGGCTGTGGCATCGAGCTCAGCTGTC 1440
Qy 1513 TTGGAGGATCTGGGCTTTTCCAGAACACAGCCGTTGAGAGCATGAATGTACCAC 1572
Db 1441 CTGGAGGATCTGGGCTTTTCCCGAAACACAGCCGTTGAAGCATGAATGTACCAT 1500
Qy 1573 GATTGGCTCTGGTGGCTCCCATCTAGGTGACACAAAGTGACACAGCAACTGTGGC 1632
Db 1501 GATTGGCTCTGGTGGCTCCCATCTAGGTGACACAAAGTGACACAGCAACTGTGGC 1560
Qy 1633 CGTCTCAGAGCTCTCGAGTGTATCTCGCCAGACACCCGTTGCGCTGGAGCTTC 1692
Db 1561 CGTCTCAGAGCTCTCAGATGTATCTCGCCAGGATCCGTTGTGCTGGAGCTTC 1620
Qy 1693 CGGCTTATCTGTGTGGCCAGCCGCGGAGCAGCCGCGGATGTTCAAGATATAGAG 1752
Db 1621 CGGCTCAGTGTGTGTGGCCATCAGAGGAGCAGCCGTTGGATGTTCAAGACATAGAG 1680
Qy 1753 TCAGCGGATCTCTCTTTGTGTCCAAAGAACCTGGGAAACATCCCGTAGTGTGAA 1812
Db 1681 TCAGCAGATCTCTCTTTGTGTCCAAAGAACCTGGGAAACATCCCGTAGTGTGAA 1740
Qy 1813 GTTCCGCTGCTACTGTGGCCAGCTGCTGCGCATGTTCCCGCAGTTCTGCTGGGCA 1872
Db 1741 GTTCCAGTGTCTACGTTGGGCGCAGTGTGCTGCGCTGCTCCCGCAGTTCTGCTGGGCA 1800
Qy 1873 TCCTGTGTGGACACAGCCAGTGGAGTGACTGCGCTCACTCCCGGAGGATGGACTA 1932
Db 1801 TCCTGTGTGGACACAGCCAGTGGAGTGACTGCGCTCACTCCCGGAGGATGGACTA 1860
Qy 1933 GAGTGTGTGATGACCCAGGGGCCATGGGGCTTTATGCTTGGAGTGTGAGAGGTGGA 1992
Db 1861 GAGTGTGTGATGACCCAGGGGCCATGGGGCTTTATGCTTGGAGTGTGAGAGGTGGA 1920
Qy 1993 GCGCCCGCTGTGGCTTATAGCTTGGTGTGGGCGAGCCAGCGGGACCCCTCAAC 2052
Db 1921 GCGCCCGCTGTGGCTTATAGCTTGGTGTGGGCGAGCCAGCGGGACCCCTCAAC 1980
Qy 2053 CGGCCCCACACCGTTGGGGCTTGGATTTGGTGGCTTCTCTGGGTGTTCTTGGACGA 2112
Db 1981 CGGCCCCACACCGTTGGGGCTTGGCTGGTGGCTTTTCTGGGTGTTCTTGGACGA 2040
Qy 2113 TCCTCTACTCTCTCTGATTGTCGCGCTCAGCAGCGTGGGCGACAGAGGAGCTTCTA 2172
Db 2041 TCCTCTACTCTCTCTGATTGTCGCGCTCAGCAGCGGCGGCGACAGAGGAGCTTCTA 2100
Qy 2173 GCTAGACACAGTGGGCTTAGATCTGGGGCTCCACCTTCTGGGACCCACAAGCTATAGT 2232
Db 2101 GCTAGACACAGTGGGCTTAGATCTGGGGCTCCACCTTCTGGGACCCACAAGCTATAGC 2160
Qy 2233 CAGGACCTCTCTCTCTCGCTGAAGATGAACGGCTGCCCTGGCCCTGGTAAAGCGG 2292
Db 2161 CAAGACCTCTCTCTCTCTCAGTGAAGATGAACGGCTGCCCTGGCCCTGGTAAAGCGG 2220
Qy 2293 GGCAGTGGTTTGGTGGCTTCCCTCCACCCCTCTCTGCTGGATTTGCCCCAAGCCAGCC 2352
Db 2221 GGCAGTGGTTTGGTGGCTTCCCTCCACCCCTCTCTGCTGGATTTGCCCCAAGCCAGCC 2280
Qy 2353 CACATCGGCTACTGGGGCGCTCTAGCCAGCTGTGATGAGACCTCCATCTAA 2406
Db 2281 CACATCGGCTACTGGGGCGCTCTAGCCAGCTGTGATGAGACCTCCATCTAA 2334

```

RESULT 7

```

US-10-170-235-32596
; Sequence 32596, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
; FILE REFERENCE: CU001380
; CURRENT APPLICATION NUMBER: US/10/170,235

```

```

; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 32596
; LENGTH: 3641
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-32596

Query Match 47.9% Score 1919.6; DB 47; Length 3641;
Best Local Similarity 77.08; Pred. No. 3.5e-251;
Matches 2699; Conservative 0; Mismatches 594; Indels 213; Gaps 21;

Qy 1 GCCAGGCCCGCCAGTAGCGGTAACTAAGTAGAGGCTGTGGAGCGCCCGCCAGCCCGCCAC 60
Db 12 GCCAGGCCAGTAGCGGTAGCCCGGGCCCTGAGCAGAGGCGGTAGCTTGGCGCCGACCCGCGG 71
Qy 61 CAGCGGAGCCAGAGATGCTTGCAGGGCGCCAGGGGCGCCCGCGGGCGCCCGCCCGCCCT 120
Db 72 CAGCGGAGCCAAAGATGCGCGGCTCTGCTGCGCGCCCGCCCGGGTCCCGGGCAGCCCT 131
Qy 121 CCGGTCTTCCCTTCCCGCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Db 132 ACAGCCTCGCCCTTCCC-----GCTACTGCTGCTGCGGCTGCTGCTGCTGCTGCTGCTG 176
Qy 181 CCGGTGTGCGCGCGCGCTCCCGCGCTCAGTCCCGCAGAACCTCGCTGCCCATCTCCGAGGCT 240
Db 177 CCGGTATCCCGCGCGCTCCCGCGCTCGGTGCGCCAGAACCTCGCTTCCAACTCTCTGAGGCT 236
Qy 241 GACTCTCTATCTCACCCTGGTTTGCAGCGCTCTCATAGTACAACTTACTGCTGCTCTCTCTG 300
Db 237 GACTCTCTATCTCACCCTGGTTTGCAGCTCCCTCACACATACAACTTACTGCTCTCTCTG 296
Qy 301 GATCTCTCCCTCCACACACTTTACGTCGGTGACGGGATAGCATCTTTCGTTTAAACCCCTC 360
Db 297 GATCTCTCCCTCCACACACTTTATGTTGGCGCCCGGACACCATCTTTCGTTTATCCCTG 356
Qy 361 CCCTTCTCTGGGGAAGACCCCGAAGATGACATGATGATGATGATGATGATGATGATGATGAT 420
Db 357 CCCTTCTCAGGGGAGAGACCCCGCAGGATTTGACTGATGATTTCTGAGGCTCACAGACAG 416
Qy 421 AACTGACAGGAAGAAGCAAGAGAGAGCAATGTACAACTTTTATCCAGATTCTCGCC 480
Db 417 AACTGTAGGAAGAAGCAAGAGAGAGCAATGTACAACTTTTGTCCAGATTCTCGCC 476
Qy 481 ATTGTCAATGCTCTCACTCTCCTCAGTCGGGACCTTCGCTTTTGTATCCGAACTCGGG 540
Db 477 ATTGCCAATGCTCTCACTCTCCTCAGTCGGGACCTTCGCTTTTGTATCCGAACTCGGG 536
Qy 541 GTTATTGATGTGTCCAGTTTCCAGAGGTTGAAAGACTTTGAGAGCGCCCGGGGAAATGT 600
Db 537 GTTATTGATGTGTCCAGTTTCCAGAGGTTGAAAGACTTTGAGAGTGGCCGGGAAATGT 596
Qy 601 CTTTTGAGCCAGCTCAACGGTCAGCAGCTGTATGCTGGGGCGCTCTCTACACCGCC 660
Db 597 CTTTTGAGCCAGCTCAACGGTCAGCAGCTGTATGCTGGGGCGCTCTCTATATGCTGCC 656
Qy 661 ACTGTGAAGAACTTCTCTGGGAGTGAAGCCCATCATCTCCCGAGCTGTGGGTGCGAGCTGAG 720
Db 657 ACTGTGAAGAACTTCTCTGGGAGTGAAGCCCATTAATACACAGAGCAGTGGGTGCGCGAG 716
Qy 721 GACTGGATTGGAACAGAGACCTTGTCTATCTCTGCTTAACTGCTCCAGCCTTTGTGCGAGCT 780
Db 717 GACTGGATTGGAACAGATACCTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 776
Qy 781 ATGGTCTCTGAGCCAGCTGAGTGGGGGATGAAGATGGAGACGATGAATCTTTTTC 840
Db 777 GTGGCTTTGAGCCAGCCGAATGGGGGATGAAGATGGAGACGACGAAATCTACTTCTTC 836
Qy 841 TTACAGGAGACCTCCCGAGTGTGGAGTCTCTATGAGCGCATCAAGTCCCAAGAGTGGCC 900
Db 837 TTACAGGAGACCTCCCGAGGATTTGACTACAGAGCGCATTAAGTCCCAAGAGTGGCC 896
Qy 901 CGAGTGTGTGGGGGACCTTTGGGGGAGGAGACCCCTTCAGCAGAGATGAGCAGCCTTT 960

```


Db 1153 CGACAGTGTGAATGTCCTTTCAGAGAACTAAACATGACTGCAACAGAGACTGCCT 1212
QY 1201 GTATTGGACCAACGAGTGGCCAGCCAGACCTGGAGAGTGCATGCCAACACATGAAG 1260
Db 1213 GTCTGGACAAATGATGTGCCACAGCCAGACCTGGAGAGTGCATCAACCAACATGAAG 1272
QY 1261 CTCACAGCTTTGGATCTCACTCTCCCTGCCAGACCGGTGTCTACCTTTATCAGAGAC 1320
Db 1273 CTCGGCACTTTGGCTCATCTCTCTCCCTGCCGTACCCGGTACTCACCTTCATCCGGAC 1332
QY 1321 CACCTCTCATGGACAGCCGGTGTCCCGGTGAGCGGCCGCCCTGCTGTGCTACTACA 1380
Db 1333 CACCCACTCATGGACAGGCGAGTGTTCAGCTGATGGCCACCCCTGCTGTGCTACTACA 1392
QY 1381 GATACAGCCTATCTCAGAGTCTGTGGCCACAGGCTGACAGCCTCTCAGGAAAGATAT 1440
Db 1393 GATACAGCCTATCTCAGAGTCTGTGGCCACAGGCTGACAGCCTCTCAGGAAAGATAT 1452
QY 1441 GAGCTCTCTACCTGGGACAGAGGATGGACACCTCCACCGGCTGTGGCATTTGAGCT 1500
Db 1453 GATGTCTCTACCTGGGACAGAGGATGGACACCTCCACCGAGCTGCGGATCGAGCT 1512
QY 1501 CAGCTCAGTGTCTTGGAGATCTGGCTTGTTCACAGAACACAGCCGTTGAGAGCTG 1560
Db 1513 CAGCTCAGGCTTCTTGAAGATCTGGCTTATTCCAGAGCCACAGCAGTTGAGAACATG 1572
QY 1561 AAATTGTACCAAGATGGCTCTGTGGCTGCCATCTCCTAGTGGTGCACAAAGTGAACAC 1620
Db 1573 AAATTGTACCAAGATGGCTCTGTGGCTGCCATCTCCTAGTGGTGCACAAAGTGAACAC 1632
QY 1621 AGCACTGTGGCGTCTCAGAGCTGCTCGAGTGTATCCTGCCCGAGGACCCGTTGTC 1680
Db 1633 ACCAAGTGTGGCGTCTCAGAGCTGCTCAGAGTGCATCCTGCCCGAGGACCCAGTCTGT 1692
QY 1681 GCCTGAGCTTCGGCTTGTATGCTGTGTGGCCACGCGCGGAGCAGCCGCGGATGGTT 1740
Db 1693 GCCTGAGCTTCGGCTTGTATGCTGTGTGGCCACGCGCGGAGCAGCCGCGGATGGTT 1752
QY 1741 CAAGATATAGAGTACAGCGATGCTCTTCTTGTGTGTCCAAAAGACCTGGAAACATCC 1800
Db 1753 CAAGATATAGAGTACAGCGATGCTCTTCTTGTGTCTTAAGAGCCTGGAAACATCC 1812
QY 1801 GTAGTGTTTAGATTCCCGTGGCTACTGTGGGCCAGCTGGTCTCGCATGTTCCCGCAGT 1860
Db 1813 GTAGTGTTTAGATTCCCGTGGCTACAGCTGGCATGTGGTCTGGCATGTTCTCAAGC 1872
QY 1861 TCTGCTGGCATCTGTGTGGCAGCAGCCAGTGGAGTCACTCGCTCACTCCCGG 1920
Db 1873 TCAGCATGGCATCTGTGTGGCAGCAGCCAGTGGAGTCACTCACTCACTCCCGG 1932
QY 1921 AGGATGGACTAGAGTGTGTGTGACCCAGGGGCTATGCTTGGCAGTGT 1980
Db 1933 CGGATGGACTAGAGTGTGTGTGACCCAGGGGCTATGCTTGGCAGTGT 1992
QY 1981 CAGGAGGTGGAGCCCGCTGTGTGTGTATAGCTTGGTGTGGGACAGCAGCG 2040
Db 1993 CAGGAGGTGGAGCCCGCTGTGTGTGTATAGCTTGGTGTGGGACAGCAGCG 2052
QY 2041 GGACCTCAACCGGGCCACACCTGTGTGGGGCTGGATGGTGTTCCTCGGT 2100
Db 2053 GATGTCCAGCGGGGCCACAC---AGTGGGGGGGAGTGGTGGCTTCTTGGGG 2109
QY 2101 GTTCTTGGACATCCTCTCACTCTCTCTGATTGGTTCGCGGTGACAGCGTGGCGACAG 2160
Db 2110 ATTCTCGAGCATCCTGACTCTCATTTGATTGGTTCGCGGTGACAGCGGCGACAG 2169
QY 2161 AGGAGCTTCTAGTAGACAAAGTGGCTTAGATCTGGGGCTCCACCTTCTGGGACC 2220
Db 2170 AGGAACTTCTGGCTAGAGACAAAGTGGGCTGGACCTGGGGCTCCACCTTCTGGGACC 2229
QY 2221 ACAAGCTATAGTCAGGACCTCCCTCTCTTCCGCTGAAGTGAACGCTGCCCTGGCC 2280

Db 2230 ACAAGCTACAGCCAGACCCCTCCCTCCCTCTCTCTGAAGATGAGCGGTTGCCGCTGGCC 2289
QY 2281 CTGGGTAAGCGGGGAGTGGTTTGGTGGCTTCCCTCCACCCCTCTCTCTGGATTCTTGC 2340
Db 2290 CTGGCCAAAGAGGGGAGTGGCTTGGTGGATTCTACACCCCTTCTCTGTTGATCTTGC 2349
QY 2341 CCAAGCCCAAGCCACATCCGGCTCACTGGGGGCGCTCTAGCCACGCTGTGATGAGACCTCC 2400
Db 2350 CCAAGCCCAAGCCACATCCGGCTCACTGGGGCTCTCTAGCCACATGTGATGAACATCC 2409
QY 2401 ATCTAAGCGGGGAAATGACTGCCAGCCATGAGCAGTCTCTGGAAC 2448
Db 2410 ATCTAGCTGGGCAATGACCACTAGTGTATAGTGATCACTGGAAC 2457

RESULT 9
US-10-305-278-65
; Sequence 65, Application US/10305278
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TEISUO
; APPLICANT: KAWAI, YURI
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: HAYASHI, KOJI
; TITLE OF INVENTION: SECRETORY PROTEIN OR MEMBRANE PROTEIN
; FILE REFERENCE: 084335/0121
; CURRENT APPLICATION NUMBER: US/10/305,278
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 05/09/611,523
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: JP 1999-194179
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: JP 2000-118775
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183766
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: 60/159,586
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,323
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 679
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 65
; LENGTH: 2971
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (103)..(2412)
US-10-305-278-65

Query Match 46.8%; Score 1875.2; DB 50; Length 2971;
Best Local Similarity 86.1%; Pred. No. 3.8e-245;
Matches 2107; Conservative 0; Mismatches 323; Indels 18; Gaps 2;

QY 1 GCCAGAGCCCGCCAGTAGCTACTAGTAGAGCTGTGTGACGCGCCCGCCCGCGCGGCT 60
Db 28 GCCAGAGCCAGTAGTCCCGGGCCCTGAGCAGAGCCGCTAGCTTGGCGGACCCGCGC 87
QY 61 CAGCGGGAGCCAGAGATGTTGCCAGGGCCGAGCGGCCCGCCCGCGCGCGCGCGCT 120
Db 88 CAGCGGGAGCCAAAGATGCCGCTCTGTGTCGCGCGCCCGCCCGCGCGCGCGCT 147
QY 121 CGGCTCTTCCCTCCCGCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Db 148 ACAGCCTCGCCCTTCCC-----GCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 192
QY 181 CCGGTGTGCGCGCGCGCTCCCGCTCAGTGGCCAGAACCTCGCTGCCCATCTCCGAGGCT 240
Db 193 CCGGTATCGCGCGCGCTCCCGCTCGGTGCGCCAGAACCTCGCTTCCCATCTCTGAGGCT 252
QY 241 GACTCCTATCTACACCGCGCTTGGACGCTCTCATAGTACAATTAATCTGCTCTCTGTTG 300

Query Match	41.8%	Score 1673.6	DB 16	Length 2893
Best Local Similarity	82.2%	Pred. No. 8.9e-218		
Matches 2012	Conservative 0	Mismatches 319	Indels 117	Gaps 3
QY	1	GCGAGCGCCGCGAGTAGCGGTACTAAGTAGAGGCTGCTGGAGCGCGCCACACCGGCAC	60	
Db	35	GCGAGCGCCAGTAGCCCCGGGGCCCTGAGCAGAGCGCGTAGCTTGCGCACACCGCGGC	94	
QY	61	CAGCGGAGCCAGAGATGTTGCCAGGGCGAGGGGCCCGCCCGCGGCCCGCCGCCGCT	120	
Db	95	CAGCGGAGCCAAAGATGCGGGCTTGCTGCGCGGCCCGCGCGGCTCCGGGACGCT	154	
QY	121	CGGCTCTTTCCTCCCGCGCGCGCTGTCGCTGCTGCTGCTGGCGATACTAAGCGCC	180	
Db	155	ACAGCTCGCCCTCC-----GCTACTGCTGCTGGCGGTGCTGAGCGGC	199	
QY	181	CGGCTGCGGCGCGCTCCCGCTCAGTGCCCGAGAACCTCGTGCGCCATCTCCGAGGCT	240	
Db	200	CGGTTATCGGCGCGCTCCCGCTCGGTGCCAGAACCTCGCTTCCAATCTCTGAGGCT	259	
QY	241	GACTCTATCTACCCGGTTTGAGCGTCTCATAGTACAATTACTCTGCTTCCTTGTC	300	
Db	260	GACTCTGTCACCGGGTTGCGAGTCCTTCACACATACAACTACTGTTCTCTTGTC	319	

QY	301	GATCCTGCCTCCACACACACTTTACGTCGGTGACCGGATAGCATCTTCGCTTTAAACCTC	360
DB	320	GATCCTGCCTCCACACACTTTATGTGGGCGCGGACACCATCTTCGCTTTATCCCTG	379
QY	361	CCCTTCTCTGGGAAACACCCGAAGATCGACTGGATGTACTGTAGACTCACAGACAG	420
DB	380	CCCTTCTCAGGGGAGACCCCGCAGAGTTGACTGGATGGTTCTTGAGGCTCACAGACAG	439
QY	421	AACTCGAGGAAGAAAGCAAGAGAGGAGGAATGTCAACAATTTATTCAGATTCGCGC	480
DB	440	AACTGTGAGAAAGCAAGAAAGAG-----	466
QY	481	ATTGTCAATGCCTCTCACTCTCTACGTCGGGCACCTTCGCTTTTGATCGAAGTGGGG	540
DB	467	-----	466
QY	541	GTTATTGATGTCCACGCTTTCCACGAGGTTGAAGACTTCAGAGCGCGGGGGAATGT	600
DB	467	-----GATGTGTCCAGGTTCCAGCAGGTTGAAGAGACTTGAGAGTGGCCCGGGGAATGT	520
QY	601	CCCTTTTGAGCAGCTCAACGGTCAGCAGCTGTAAATGGCTGGGGGCGTCTCTACACCGCC	660
DB	521	CCCTTTTGAGCCAGCTCAGCGTCAGCAGCTGTAAATGGCTGGGGGGTCTCTATGCTGCC	580
QY	661	ACTGTGAAGAACTTCTCGGGAGCTGAGCCCATCATCTCCGAGCTGTGGGTGAGCTGAG	720
DB	581	ACTGTGAAAACACTCCCTGGGAGCGGAGCCAAATATCACCAGCAGCTGGGTCTGCGCGAG	640
QY	721	GACTGGATTCCAGACAGACCTTCTCATCTCGCTTAATCTCCAGCCTTGTCTCGACGT	780
DB	641	GACTGGATTCCGACAGATCTTGCCTTCTCGCTGAACGCCCCAGCGCTTGTGCGAGCC	700
QY	781	ATGCTCTGAGCCACGCTGAGTGGGGGATGAAGATGGAGCAGATGAATCTTTTTC	840
DB	701	GTGGCCTTGAGCCAGCCGAAATGGGGGATGAAGATGGAGACCAAGAAATCTACTCTTC	760
QY	841	TTCACGGAGACCTCCCGAGTGTGGACTCCTATGAGCGCATCAAGTCCCAAGATGGCC	900
DB	761	TTTACGGAGACTTCCCGAGCATTTGACTCATACGAGCATTAAGTCCCAAGTGGCC	820
QY	901	CGAGTGTGCGGGGACCTTGGGGCAGGAAGACCTTCAGCAGAGATGGACGAGTTT	960
DB	821	CGTGTGTGCGGGGACCTCGGGGGCCGAAAGACCTCCAGCAGAGATGGACGAGTTT	880
QY	961	CTGAAGGCTGACCTGTCTGCCACAGGSCCGAGCATGGCGGGCTCTCGGGTCTTCGAC	1020
DB	881	TTGAAAGCTGACCTGTCTCTCCAGGCTGAGCATGGCGGGCTCCAGTGTCTCTGCAG	940
QY	1021	GCTATGCGAGAGCTTCGGCTCAGCCTGGAGCGGAAACCCCATCTTTTATGGGATCTTT	1080
DB	941	GATGTTGCTGTGCTTCGACCTGAGCTTGGGGCAGGGACTCCCATCTTTTATGCACTTT	1000
QY	1081	TCCTCCAGTGGGAAGAGCTGCCATCTCTGCTGTGTGTGCTCTCGACCCCAAGACATC	1140
DB	1001	TCCTCCAGTGGGAGGGGCTACTATCTCTGCTGTCTGTGCTCTCGACCCACCAAGACA	1060
QY	1141	CGGCGAGTGTGAATGTCTCTTTAGAGAGCTAAACATGACATGCAACAGGGGACCTGC	1200
DB	1061	CGGACAGTGTGAATGTCTCTTCAGAGAACTAAACATGACATGCAACAGGAGCTGCCT	1120
QY	1201	GTATGACAACAGAGTGTCCCGCAGCCAGACCTGGAGAGTGCATCGCCAAACAACATGAAG	1260
DB	1121	GTGCTGGCAATGATGTGCCCGAGCCAGACCTGGAGATGTCATCACCACACATGAAG	1180
QY	1261	CTCCAGAGTTTGGATCTCACTCTCCCTGCCAGACCGGGTGTCTACCTTTATCAGAGAC	1320
DB	1181	CTCCGGACATTTGGCTCATCTCTCTCCCTGACCGCGTACTCACTCTCATCTCCGGGAC	1240
QY	1321	CACCTCTCATGACAGGCCCGTGTCCCGGCTGAGGGCGCCCTCTGCTGGTCACTACA	1380
DB	1241	CACCCACTCATGACAGGCGAGTGTTCAGCTGTATGGGCAACCCCTGCTGTGTCACATA	1300
QY	1381	GATACAGCCTTATCTCAGAGTCGTGGGCCACAGGGTGACCGCCTCTCAGGGAAAGAAAT	1440


```

1001 TCTTCCAGTGGGAGGGGCTACTATCTCTGTGTGTGTCTTCCGACCACAAGACATT 1060
1141 CGGCAGTGTGTAATGTCCTCTTAGAGAGCTAAACATGACTCAACAGGGGACTGCT 1200
1061 CGGACAGTGTGTAATGTCCTCTCAGAGACTAAACATGACTCAACAGAGGACTGCT 1120
1201 GTCATGGACAAACAGAGGTGCCCCAGACCTGGAGAGTGCATCGCCAAACAATGAAG 1260
1121 GTCTGGACAATGATGTCGCCAGCCAGACCTGGAGAGTGCATCACCAACAATGAAG 1180
1261 CTCAGAGAGTTGGATCTCACTCTCCCTGCCAGACCCGCTGCTCACCTTTATCAGAGAC 1320
1181 CTCGGCAGTTTGGCTCATCTCTCCCTGCTGACCGCTACTCACCTTCATCCGGGAC 1240
1321 CACCTCTCATGGACAGGCCGCTGTCCCGCTGAGGCCGCCGCCCTCTCTGTCTACTACA 1380
1241 CACCACTCATGGACAGGCCAGTGTTCACAGCTGATGCCACCCCTGCTGTCTACTACA 1300
1381 GATACAGCTATCTCAGAGTCTGGCCACAGGGGTGACAGCCTCTCAGGGAAGAATAT 1440
1301 GATACAGCTATCTCAGAGTCTGGCCACAGGGGTGACAGCCTCTCAGGGAAGAATAT 1360
1441 GACGTGCTTACCTGGGACAGAGATGGACACCTCCACCCGGCTGTGCGCATTTGGAGCT 1500
1361 GATGTGCTTACCTGGGACAGAGATGGACACCTCCACCCAGAGTGTGCGCATTTGGAGCT 1420
1501 CAGCTCAGTGTCTGGAGGATCTGGCCTTGTCCAGAAACACACAGCCGTTGAGAGCATG 1560
1421 CAGCTCAGCTTCTTGAGATCTGGCCTTATCCAGAGCCACAGCCAGTTGAGNACATG 1480
1561 AAATGTATACACAGTATGGCTCCTGGTGGCTCCCATCTGAGGTGACAAAGTGAACACC 1620
1481 AAATGTATACACAGTGTGCTCTGGTGGCTCCCATCTGAGGTGACAAAGTGAATACA 1540
1621 AGCAACTGTGGCGCTCTCAGAGTGTCTCGAGGTATCTCTGGCCAGGACCCCGTGGC 1680
1541 ACCAACTGTGGCGCTCTCAGAGTGTCTCGAGGTATCTCTGGCCAGGACCCCGTGGT 1600
1681 GCCTGGAGCTTCCGGCTTGTGATGTGTGGCCACAGCCGGCGAGACACCGGGGATGTT 1740
1601 GCCTGGAGCTTCCGGCTGATGTGTGGCCATGCTGGCCATGCGGGGAGCACCAGGGTGTGTC 1660
1741 CAAGATATAGAGTACAGCGGATGTCTCTTTGTGTGTCCAAAAGAACCTTGGAGAACATCC 1800
1661 CAAGACATAGAGTACAGAGATGTCTCTCTTTGTGTCTTAAAGAGCCTTGGAGAACGTCA 1720
1801 GTAGTGTGTAAGTTCCGGTGTCTACTGTGGCCACAGTGGTCTCCCATGTTCCCCAGT 1860
1721 GTAGTGTGTAAGTTCCCGTGTCTACAGCTGCGCATGTGGTCTTGCCATGTTCTCCAAGC 1780
1861 TCTGCTGGGATCCTGTGTGTGGCCACAGCCAGTGGAGTGAATGCTCACTCCCGGG 1920
1781 TCAGATGGGATCCTGTGTGTGGCCACAGCCAGTGGAGTGAATGCTCACTCCCGGG 1840
1921 AGGATGAGCTAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1980
1841 CGGATGAGCTAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1900
1981 CAGAGGCTGAGCGCCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2040
1901 CAGAGGCTGAGCGCCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1960
2041 GGACCTCAAAACCGGGCCACACCTTGTGGGGCTGGATTTGGTGGCTTTCTCTGGGT 2100
1961 GATGCTCGAGCGCGGGCCACAC --- AGTGGGGGGGAGTGGTGGCTTTCTTCTTGGGG 2017
2101 GTTCTTGAGCATCCCTCACTCTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2160
2017 GTTCTTGAGCATCCCTCACTCTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2077
2161 AGGAGCTTCTAGCTAGAGCAAGGTGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2220
2077 AGGAACTTCTGCTAGAGCAAGGTGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2137
```


Qy	2221	ACAAGCTATAGTCAGGACCCCTCCCTCTCCTTTCGCCCTGAAGATGAACGGGTGCCCGCTGGCC	2280
Db	2138	ACAAGCTACAGCCAAGACCCTCCCTCCCCTCTCCTGAAGATGAGCGTTGCCTGGCC	2197
Qy	2281	CTGGGTAAAGCGGGCAGTGTTTTGGTGGCTTCCCTCCACCCCTCTCTGCTGGATTCTTGC	2340
Db	2198	CTGGCCAAGAGGGCAGTGGCTTTGGTGGATTCTCACCACTCTCTGCTTGATCCTTGC	2257
Qy	2341	CAAAGCCCCACCATCCGGCTCACTGGGGCGCCTCTAGCCACGTTGATGAGACCTCC	2400
Db	2258	CRAAGCCCAACCCACATTCGGCTTAACCTGGGGCTCTCTAGCCACATGTGATGAACATCC	2317
Qy	2401	ATCTAAGCGGGGAAAATGACTGCCAGCCATGAGCAGTCTCTTGGAAC	2448
Db	2318	ATCTAGAGTGGGCAATGACCACCTAGTGATAAGTGATCACTGGAAC	2365

```

RESULT 13
US-10-036-657-41
; Sequence 41, Application US/10036657
; GENERAL INFORMATION:
; APPLICANT: Earl Francis Albone, et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES
; FILE REFERENCE: GP-70778B-CI
; CURRENT APPLICATION NUMBER: US/10/036,657
; CURRENT FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 41
; LENGTH: 2893
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-10-036-657-41

```

Query Match	41.98;	Score	1673.6;	DB	44;	Length	2893;
Best Local Similarity	82.24;	Pred.	No. 8.9e-218;				
Matches 2012;	Conservative	0;	Mismatches	117;	Indels	117;	Gaps
3;							
Qy	1	GCGAGGCGCGCGAGTAGCGGTACTAAGTAGAGGCTGCTGGACGGCCGCCACCCCGGCAC	60				
Db	35	GCGAGGCCAGTAGCCCCGGGCGCTTCAGCAGAGGCCGTAGCTTGGCCGCCACCCCGGGC	94				
Qy	61	CAGGCGAGCCAGAGATGCTTCCAGGGCGGAGCGGCCGCCGCCGGGCCCGCGCCGCT	120				
Db	95	CAGGCGGAGCAAGATGCGGCGCTCTGCTGCGCGGCCGCCGCCCGGCTCCGGGACGCT	154				
Qy	121	CGGCTCTTCCCTTCCCGCGCGCGCTGCTCGCTGCTGCTGCTGCGGTACTAAGCGGC	180				
Db	155	ACAGCTCGGCCCTCC-----GCTACTGCTGCGGCTGTGAGCGGC	199				
Qy	181	CGGCTGTGGCGCGCTCCCGCTCAGTGCCAGAACCTGCTGCGCCATCTCCGAGGCT	240				
Db	200	CCGCTATCCGCGCGCTCCCGCTCGGTGCCAGAACCTCGCTTCCAATCTCTGAGGCT	259				
Qy	241	GACTCTATCTCACCGGTTTGACGGCTCTCATACGTACAAATTACTCTGCTCTCCTTG	300				
Db	260	GATCTCTGTCTACCGGTTGCGAGTCCCTCACATACAAATTACTGTGTTCTCCTTG	319				
Qy	301	GATCCTGCCTCCACACACTTTACGTCGGTGCAGGGTAGCATCTTCGTTTAAACCTC	360				
Db	320	GATCCTGCCTCCACACACTTTATTTGGCGCCGGACACCATCTTCGCTTATCCCTG	379				
Qy	361	CCCTTCTCTGGGAAGACCCCGAAGATGCGTGTGATGTGCTGAGACTCACAGAC	420				
Db	380	CCCTTCTCAGGGGAGAGACCCCGAGGATTTGACTGGATGGTTCTCGAGGCTCACAGAC	439				
Qy	421	AACTGAGGAAGAAAGCAGAAAGAGGAGATGTCACAATTTTATCCAGATTCTCGCC	480				
Db	440	AACTGTAGGAAGAAAGCAAGAAG-----	466				
Qy	481	ATTGTCAATGCCTCTCACCTCTCAGGTGCGGACCTTCGCTTTTGTATCCGAAGTCGGG	540				

Db	467	-----	466
Qy	541	GTATTGATGTCCTCCAGCTTTCCAGCAGGTTGAAAGACTTTGAGACGGCCGGGGGAAATGT	600
Db	467	-----GATGTCTCCAGGTTCCAGCAGGTTGAAAGACTTGAGAGTGCCGGGGGAAATGT	520
Qy	601	CTTTTTCAGCCAGCTCAACGGTCAGCAGCTGTAAATGGCTGGGGGGCTCTCTACACCGCC	660
Db	521	CTTTTTCAGCCAGCTCAGCGGTCAGCAGCTGTAAATGGCTGGGGGGCTCTCTATGCTGCC	580
Qy	661	ACTGTGAAGAACTTCTGTGGGACTTGAGCCCATCATCTCCGAGCTCTGGGTTCAGCTGAG	720
Db	581	ACTGTGAAAAACTACTCTGGGAGCGGACCCAAATTATCACACAGCAGTGGGTCTGCCGAG	640
Qy	721	GACTGGATTCCGAACAGAGACCTTGTTCATCTGCTTAAATGCTCCAGCCCTTTGTCGAGCT	780
Db	641	GACTGGATTTCGACAGATACCTTTCCTTCCTGCTGAAGCCCCAGCCTTTGTGTCGAGCC	700
Qy	781	ATGGTCTGAGCCACTGAGTCGGGGGATGAAGATGGAGCAGTCAAAATCTTTTTC	840
Db	701	GTGGCTTGAGCCAGCCGAATGGGGGATGAAGATGGAGACCGCAAAATCTACTTCTTC	760
Qy	841	TTACAGGAGACCTCCCGAGTGTGGACTTCCTATGAGCGCATCAAGTCCCAAGAGTGGCC	900
Db	761	TTTACGAGACTTCCCGAGCAATTGACTATACGAGCGCAATTAAGTCCACACGGGTGGCC	820
Qy	901	CGAGTGTGCGGGGACCTTGGGGGCGAGAAACACCTTTCAGCAGAGATGGACGACGTTT	960
Db	821	CGTGTGTGCGGGGACCTTGGGGGCGGAACCCCTCCAGCAGAGATGGAGCAGCTTT	880
Qy	961	CTGAAGGCTGACCTGTGTGCCAGGGCCCGAGCATGGCCGGGCTTCGCGGTTCTGCAG	1020
Db	881	TTGAAAGCTGACCTGCTGTCTCAGGGGCTGACCATGGCCGGGCTCCAGTGTCTCTCAG	940
Qy	1021	GCTATGGCAGAGCTTGGCCCTCAGCCTGGAGCGGGAAACCCCAATCTTTTATGGGATCTTT	1080
Db	941	GATGTTGCTGTGTTGCACCTGAGCTTGGGGCAGGACTCCCACTCTTTTATGGCATCTTT	1000
Qy	1081	TCCTCCAGTGGGAAGAGCTGCCATCTCTGCTGTGTGTGCCCTTCGGACCCCAAGACATC	1140
Db	1001	TCYTCCAGTGGAGGGGCTACTATCTCTGCTGTCTGTGCCCTTCGACCAACAGACATT	1060
Qy	1141	CGGCGAGTGTGTAATGGTCCCTTTAGAGAGCTAAAACATGACTGCAACAGGGGACTGCT	1200
Db	1061	CGGACAGTGTGTAATGGTCCCTTTCAGAGAACTAAAACATGACTGCAACAGAGGACTGCT	1120
Qy	1201	GTCAATGCACAAACAGGTGCCCGCCAGACCTGGAGTGCATGCCCAACACATGAAG	1260
Db	1121	GTGCTGGACAANTGATGTGCCCGCCAGACCTTGGAGGTGCATACCAACACATGAAG	1180
Qy	1261	CTCCAGCAGTTTGGATCTCACTCTCCCTGCCAGACCGGTGCTCACTCTTATCAGAGAC	1320
Db	1181	CTCCGGCAGTTTGGCTCATCTCTCTCCCTGACCGGTACTCACTCTCATCCGGGAC	1240
Qy	1321	CACCTCTCATGACAGCGCCGTTTCCGGGCTAGCGCGGCCCTGCTGCTCACTACA	1380
Db	1241	CACCCACTCATGACAGCGCCAGTGTTCACGTGTATGGCCACCCCTGCTGCTCACTACA	1300
Qy	1381	GATACAGCCTTATCTCAGAGTGTGGGCCACAGGGTGACCGCTCTCAGGGAAAGAAATAT	1440
Db	1301	GATACAGCCTTATCTCAGAGTGTGGGCCACAGGGTGACCGCTCTCAGGGAAAGAGTAT	1360
Qy	1441	GAGTGTCTTACCTGGGACAGAGATGNACACTCCACGGGCTGTGGCATTTGAGAGCT	1500
Db	1361	GATGTGCTCTTACCTGGGACAGAGATGNACACTCCACCGAGCTGTGGGATTCGGAGCT	1420
Qy	1501	CAGCTCAGTCTTTGGAGATCTGGCTTGTTCACAGAACACACCGCGTTTGAGAGCATG	1560
Db	1421	CAGCTCAGCGTTCTTGAGAGATCTGGCTTATTTCCAGAGCCACAGCGCTTGAGACATG	1480
Qy	1561	AAATTGTACCAAGATTGGCTCTCTGGTGGGCTCCCACTACTGAGGTGACACAAGTGAACACC	1620
Db	1481	AAATTGTACCAAGCTGGTCTCTGGTGGGCTCCCGTACTGAGGTGACACAAGTGAATACA	1540

Qy	1021	GCTATGCGAGAGCTTCCGCTCAGCCTGGAGCGGGAACCCCACTCTTTTATGGGATCTTT	108
Db	941	GATGTTCTGTGCTTCACCTTGAGCTTTGGGCGAGGACTCCCACTCTTTTATGGCATCTTT	1000
Qy	1081	TCCTCCAGTGGGAAGAGCTGCCATCTCTGCTGTGTGTGCTGTCCGACCCCAAGACATC	1140
Db	1001	TCCTCCAGTGGGAGGGGCTACTATCTCTGCTGTCTGTGCTTCCGACCCACAGACATTT	1060
Qy	1141	CGGSCAGTGTGAAATGGTTCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCCT	1200
Db	1061	CGGACAGTGTGTAATGTGCTCTTCAGAGAACTAAACATGACTGCAACAGAGGACTGCCT	1120
Qy	1201	GTCATGACAAACAGAGTGGCCAGCCAGCAGCTGGAGAGTGATCGGCAACAAATGAAG	1260
Db	1121	GTCGTGACAAATGATGTGCCACAGCCAGACCTGGAGAGTGATCAACCAACAACTGAAG	1180
Qy	1261	CTCCAGAGATTTGGATCCTCACTCTCCCTGCCACACCGGTGCTCACTTTTATCAGAGAC	1320
Db	1181	CTCCGGACATTTGGCTCATCTCTCTCCCTGGCTGACCGGTACTCACTTCACTCCGGAC	1240
Qy	1321	CACCTCTCATGGACAGCCCGTGTTCGCCGCTGACGCCGCCGCCCTGCTGTGTCACATACA	1380
Db	1241	CACCCACTCATGGACAGGCCAGTGTTCAGCTGATGGCCACCCCTGCTGTGTCACATACA	1300
Qy	1381	GATACAGCTATCTACAGTCTGTGGCCACAGGTGTACCAGCTCTCAGGGAAGAATAT	1440
Db	1301	GATACAGCTATCTACAGTCTGTGGCCACAGGTGTACCAGCTCTCAGGGAAGAAGTAT	1360
Qy	1441	GACGTGCTACCTGGGACAGAGATGGACACCTCCACCGGGGTGTGCGCATTTGGAGCT	1500
Db	1361	GATGTGCTTACCTGGGACAGAGATGGACACCTCCACCGAGCAGTGTGCGGATCGGAGCT	1420
Qy	1501	CAGCTCAGTGTCTTTGGAGATCTTGGCCTTGTTCAGAACACACAGCCGGTTGAGAGCATG	1560
Db	1421	CAGCTCAGCGTCTTTGAAGATCTTGGCCTATTTCAGAGCCACAGCCAGTTGAGAACATG	1480
Qy	1561	AAATTGTACCAAGATTTGGCTCCTGGTGGGCTCCCATACTGAGGTGACACAAGTGAACACC	1620
Db	1481	AAATTGTACCAAGTGTGGCTCCTGGTGGGCTCCCGTACTGAGGTGACACAAGTGAATACA	1540
Qy	1621	AGCAACTGTGGCGCTCTCCAGAGCTGTCCGAGTGTATCTCGGCCACAGGACCCGTGTGC	1680
Db	1541	ACCAACTGTGGCGCTCTCCAGAGCTGTCCAGAGTGCACTTGGCCAGGACCCAGTCTGT	1600
Qy	1681	GCCTGGAGCTTCCGGCTTGATGCTTGTGTGGCCACCGCGGAGCAGCACCGCGGATGGTT	1740
Db	1601	GCCTGGAGCTTCCGGCTGGATGACTGTGTGCCCATGCCGGGAGCACCGAGGTTGGTC	1660
Qy	1741	CAAGATATAGAGTCAGCGGATGTCTCTTCTTTGTGTCCAAAGAACCTGGAGACATCCC	1800
Db	1661	CAAGACATATAGATCAGCAGATGTCTCCCTTCTTTGTGTCTTAAAGAGCTGGAGAACGTCCA	1720
Qy	1801	GTACTGTTGAAGTTCGGTGGCTACTGTGTGGCCAGCTGTGCTGCCATGTTCCCCAGT	1860
Db	1721	GTAAGTGTGAAGTTCGGTGGCTTACAGCTCGGCATGTGGTCTTGCCATGTTCTCCAAGC	1780
Qy	1861	TCTGCCTGGGCATCCTGTGTGGCACAGCCCACTGGAGTGACTGCGCTCACTCCCGGG	1920
Db	1781	TCAGCATGGGCATCCTGTGTGGCACAGCCAGCCAGTGGAGTGACTCACTCCCGCGG	1840
Qy	1921	AGGATGAGCTATAGAGTGGTGGTGACCCCAAGGGGCATGGGGGCTTATGCTTGGAGTGT	1980
Db	1841	CGGATGGACTGGAGTGGTGGTGACCCAGGGGCCATGGGGCTTATGCTGTGAATGT	1900
Qy	1981	CAGAGAGGTGAGCCGCCCGCTGGTGGCTGCTTATAGCTTGGTGTGGGCGACCGCGG	2040
Db	1901	CAGAGAGGTGGGGCAGCCCATGTGTGTAGCAGCTTACAGCTTGGTATGGGGCAGCCAGGA	1960
Qy	2041	GGACCTCTAAACCGGGCCACACCTGTGTGGGGCTGGATGGTGTGGCTTCTCTCTGGT	2100
Db	1961	GATGCTCGACCGGGGCCACAC---AGTGGGGGGGACTGGCTGGCTTCTCTTGGGG	2020
Qy	2101	GTTCTTGCAGCATCCCTCACTCTCCTCTCTGATTTGGTCGCCGTCAAGCAGCTGTGGCAGAC	2160

Result No.	Score	Query Match	Length	DB	ID	Description	
1	218	9.4	2433	4	US-09-300-958A-24	Sequence 24, Appl	
2	215.2	9.2	4157	4	US-08-556-422A-1	Sequence 1, Appl	
3	94.2	4.0	3524	4	US-09-077-940A-3	Sequence 3, Appl	
4	79.2	3.4	3892	4	US-09-077-940A-1	Sequence 1, Appl	
5	66.6	2.9	2854	1	US-08-121-713D-57	Sequence 57, Appl	
6	66.6	2.9	2854	1	US-08-835-268-57	Sequence 57, Appl	
7	66.6	2.9	2854	2	US-09-060-692-57	Sequence 57, Appl	
8	66.6	2.9	2854	3	US-08-833-391-57	Sequence 57, Appl	
9	66.6	2.9	2854	4	US-09-060-610-57	Sequence 57, Appl	
10	66.6	2.9	2854	5	PCR-US94-10151A-57	Sequence 57, Appl	
11	60.6	2.6	2601	1	US-08-121-713D-53	Sequence 53, Appl	
12	60.6	2.6	2601	1	US-08-835-268-53	Sequence 53, Appl	
13	60.6	2.6	2601	2	US-09-060-692-53	Sequence 53, Appl	
14	60.6	2.6	2601	3	US-08-833-391-53	Sequence 53, Appl	
15	60.6	2.6	2601	4	US-09-060-610-53	Sequence 53, Appl	
16	60.6	2.6	2601	5	PCR-US94-10151A-53	Sequence 53, Appl	
17	57.6	2.5	2898	4	US-09-308-179B-2	Sequence 2, Appl	
18	56.8	2.4	7218	1	US-08-232-463-14	Sequence 14, Appl	
19	51.8	2.2	2504	1	US-08-121-713D-63	Sequence 63, Appl	
20	51.8	2.2	2504	1	US-08-835-268-63	Sequence 63, Appl	
21	51.8	2.2	2504	2	US-09-060-692-63	Sequence 63, Appl	
22	51.8	2.2	2504	3	US-08-833-391-63	Sequence 63, Appl	
23	51.8	2.2	2504	4	US-09-060-610-63	Sequence 63, Appl	
24	51.8	2.2	2504	5	PCR-US94-10151A-63	Sequence 63, Appl	
25	51	2.2	3560	1	US-08-121-713D-59	Sequence 59, Appl	
26	51	2.2	3560	1	US-08-835-268-59	Sequence 59, Appl	
27	51	2.2	3560	2	US-09-060-692-59	Sequence 59, Appl	

```
QY 472 GATGTGTCAGATTTCCAGCAGGTTGAA-----AGACTTGAGACGGC 513
D 473 GATGTGTCAGATTTCCAGCAGGTTGAA-----AGACTTGAGACGGC 514
D 474 GATGTGTCAGATTTCCAGCAGGTTGAA-----AGACTTGAGACGGC 515
QY 514 CGGGGAAATGCTTTTTCAGCAGGTTGAA-----AGACTTGAGACGGC 573
D 515 CGGGGAAATGCTTTTTCAGCAGGTTGAA-----AGACTTGAGACGGC 574
D 516 CGGGGAAATGCTTTTTCAGCAGGTTGAA-----AGACTTGAGACGGC 575
QY 382 AAGGTCATGTTGCTTTCAGCAGGTTGAA-----AGACTTGAGACGGC 441
D 383 AAGGTCATGTTGCTTTCAGCAGGTTGAA-----AGACTTGAGACGGC 442
D 384 AAGGTCATGTTGCTTTCAGCAGGTTGAA-----AGACTTGAGACGGC 443
QY 574 CTCTACACCGCCTGCTGAGACCTTCCCTGGGAGCTGAGCCATCATCTCCGAGCTG 633
D 575 CTCTACACCGCCTGCTGAGACCTTCCCTGGGAGCTGAGCCATCATCTCCGAGCTG 634
D 576 CTCTACACCGCCTGCTGAGACCTTCCCTGGGAGCTGAGCCATCATCTCCGAGCTG 635
QY 442 CTCTACACCGCCTGCTGAGACCTTCCCTGGGAGCTGAGCCATCATCTCCGAGCTG 495
D 443 CTCTACACCGCCTGCTGAGACCTTCCCTGGGAGCTGAGCCATCATCTCCGAGCTG 496
D 444 CTCTACACCGCCTGCTGAGACCTTCCCTGGGAGCTGAGCCATCATCTCCGAGCTG 497
QY 634 GGTGAGCTGAGACCTGAGTTCAGACAGACCTTGTCTATCTCTGCTTAACTCCAGCC 693
D 635 GGTGAGCTGAGACCTGAGTTCAGACAGACCTTGTCTATCTCTGCTTAACTCCAGCC 694
D 636 GGTGAGCTGAGACCTGAGTTCAGACAGACCTTGTCTATCTCTGCTTAACTCCAGCC 695
QY 496 AGCCAGAGTTCGCGCCCAACAGAGCTGAGAGCTCCCTCAACTGGGTCAAGACCTGCC 555
D 497 AGCCAGAGTTCGCGCCCAACAGAGCTGAGAGCTCCCTCAACTGGGTCAAGACCTGCC 556
D 498 AGCCAGAGTTCGCGCCCAACAGAGCTGAGAGCTCCCTCAACTGGGTCAAGACCTGCC 557
QY 694 TTTGTCGAGCTATGCTGAGCCAGCTGAGTGGGGAGTGAAGATGGAGACAGTAA 753
D 695 TTTGTCGAGCTATGCTGAGCCAGCTGAGTGGGGAGTGAAGATGGAGACAGTAA 754
D 696 TTTGTCGAGCTATGCTGAGCCAGCTGAGTGGGGAGTGAAGATGGAGACAGTAA 755
QY 556 TTTGTCGAGCTATGCTGAGCCAGCTGAGTGGGGAGTGAAGATGGAGACAGTAA 615
D 557 TTTGTCGAGCTATGCTGAGCCAGCTGAGTGGGGAGTGAAGATGGAGACAGTAA 616
D 558 TTTGTCGAGCTATGCTGAGCCAGCTGAGTGGGGAGTGAAGATGGAGACAGTAA 617
QY 754 ATCTTTTTCCTTTCAGGAGACCTCCCGAGTGTGGAGCTTCTATGAGCGCATCAAGTC 813
D 755 ATCTTTTTCCTTTCAGGAGACCTCCCGAGTGTGGAGCTTCTATGAGCGCATCAAGTC 814
D 756 ATCTTTTTCCTTTCAGGAGACCTCCCGAGTGTGGAGCTTCTATGAGCGCATCAAGTC 815
QY 616 ATCTACTTCTTTCAGGAGACCTCCCGAGTGTGGAGCTTCTATGAGCGCATCAAGTC 675
D 617 ATCTACTTCTTTCAGGAGACCTCCCGAGTGTGGAGCTTCTATGAGCGCATCAAGTC 676
D 618 ATCTACTTCTTTCAGGAGACCTCCCGAGTGTGGAGCTTCTATGAGCGCATCAAGTC 677
QY 814 CCAAGAGTGGCGGAGTGTGCGGGGAGCTTGGGGAGGAGGAGACCTTCCAGCAGAGA 873
D 815 CCAAGAGTGGCGGAGTGTGCGGGGAGCTTGGGGAGGAGGAGACCTTCCAGCAGAGA 874
D 816 CCAAGAGTGGCGGAGTGTGCGGGGAGCTTGGGGAGGAGGAGACCTTCCAGCAGAGA 875
QY 676 TCCGAGTGTCCGAGCTGTGAAGGCGATGAGGTGGAGAGCGGTGTGAGCAACGC 735
D 677 TCCGAGTGTCCGAGCTGTGAAGGCGATGAGGTGGAGAGCGGTGTGAGCAACGC 736
D 678 TCCGAGTGTCCGAGCTGTGAAGGCGATGAGGTGGAGAGCGGTGTGAGCAACGC 737
QY 874 TGGAGACCTTTCAGGAGCTGAGTGTGCGGAGGAGCTGAGTGGGGAGGAGGAGGAG 933
D 875 TGGAGACCTTTCAGGAGCTGAGTGTGCGGAGGAGCTGAGTGGGGAGGAGGAGGAG 934
D 876 TGGAGACCTTTCAGGAGCTGAGTGTGCGGAGGAGCTGAGTGGGGAGGAGGAGGAG 935
QY 736 TGGAGACCTTTCAGGAGCTGAGTGTGCGGAGGAGCTGAGTGGGGAGGAGGAGGAG 795
D 737 TGGAGACCTTTCAGGAGCTGAGTGTGCGGAGGAGCTGAGTGGGGAGGAGGAGGAG 796
D 738 TGGAGACCTTTCAGGAGCTGAGTGTGCGGAGGAGCTGAGTGGGGAGGAGGAGGAG 797
QY 934 GGGGTCTTCAGGAGCTGAGTGTGCGGAGGAGCTGAGTGGGGAGGAGGAGGAGGAG 990
D 935 GGGGTCTTCAGGAGCTGAGTGTGCGGAGGAGCTGAGTGGGGAGGAGGAGGAGGAG 991
D 936 GGGGTCTTCAGGAGCTGAGTGTGCGGAGGAGCTGAGTGGGGAGGAGGAGGAGGAG 992
QY 796 AAGCTGCTACAGATGCTTTCAGGAGCTGAGTGTGCGGAGGAGCTGAGTGGGGAGGAG 855
D 797 AAGCTGCTACAGATGCTTTCAGGAGCTGAGTGTGCGGAGGAGCTGAGTGGGGAGGAG 856
D 798 AAGCTGCTACAGATGCTTTCAGGAGCTGAGTGTGCGGAGGAGCTGAGTGGGGAGGAG 857
QY 991 TTTTATGAGGAGCTTTCAGGAGCTGAGTGTGCGGAGGAGCTGAGTGGGGAGGAGGAG 1044
D 992 TTTTATGAGGAGCTTTCAGGAGCTGAGTGTGCGGAGGAGCTGAGTGGGGAGGAGGAG 1045
D 993 TTTTATGAGGAGCTTTCAGGAGCTGAGTGTGCGGAGGAGCTGAGTGGGGAGGAGGAG 1046
QY 856 TCTATCGGGGTCTTTCAGGAGCTGAGTGTGCGGAGGAGCTGAGTGGGGAGGAGGAG 915
D 857 TCTATCGGGGTCTTTCAGGAGCTGAGTGTGCGGAGGAGCTGAGTGGGGAGGAGGAG 916
D 858 TCTATCGGGGTCTTTCAGGAGCTGAGTGTGCGGAGGAGCTGAGTGGGGAGGAGGAG 917
QY 1045 GCCTTCGAGACCTGAGATGAGTGTGCGGAGGAGCTGAGTGGGGAGGAGGAGGAGGAG 1104
D 1046 GCCTTCGAGACCTGAGATGAGTGTGCGGAGGAGCTGAGTGGGGAGGAGGAGGAGGAG 1105
D 1047 GCCTTCGAGACCTGAGATGAGTGTGCGGAGGAGCTGAGTGGGGAGGAGGAGGAGGAG 1106
QY 916 GTCTTCACCATGAGTGTGAGAGGCTTTCAGGAGGCTTTCAGGAGGCTTTCAGGAGG 975
D 917 GTCTTCACCATGAGTGTGAGAGGCTTTCAGGAGGCTTTCAGGAGGCTTTCAGGAGG 976
D 918 GTCTTCACCATGAGTGTGAGAGGCTTTCAGGAGGCTTTCAGGAGGCTTTCAGGAGG 977
QY 1105 GACTGCAACAGGGAGCTGCTGATGAGACAGAGGCTGAGTGGGGAGGAGGAGGAGGAG 1164
D 1106 GACTGCAACAGGGAGCTGCTGATGAGACAGAGGCTGAGTGGGGAGGAGGAGGAGGAG 1165
D 1107 GACTGCAACAGGGAGCTGCTGATGAGACAGAGGCTGAGTGGGGAGGAGGAGGAGGAG 1166
QY 976 GAGACACAGCTG---GTATACGAGACCCAGGCTGAGTGGGGAGGAGGAGGAGGAG 1032
D 977 GAGACACAGCTG---GTATACGAGACCCAGGCTGAGTGGGGAGGAGGAGGAGGAG 1033
D 978 GAGACACAGCTG---GTATACGAGACCCAGGCTGAGTGGGGAGGAGGAGGAGGAG 1034
QY 1165 TGCATCGCCCAACATGAGCTCCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1224
D 1166 TGCATCGCCCAACATGAGCTCCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1225
D 1167 TGCATCGCCCAACATGAGCTCCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1226
QY 1033 TGCATTCACCATGAGTGTGAGAGGCTTTCAGGAGGCTTTCAGGAGGCTTTCAGGAGG 1092
D 1034 TGCATTCACCATGAGTGTGAGAGGCTTTCAGGAGGCTTTCAGGAGGCTTTCAGGAGG 1093
D 1035 TGCATTCACCATGAGTGTGAGAGGCTTTCAGGAGGCTTTCAGGAGGCTTTCAGGAGG 1094
QY 1225 GTGCTCACCTTATCAGAGACCCCTTCATGAGGAGGCTGAGTGGGGAGGAGGAGGAG 1284
D 1226 GTGCTCACCTTATCAGAGACCCCTTCATGAGGAGGCTGAGTGGGGAGGAGGAGGAG 1285
D 1227 GTGCTCACCTTATCAGAGACCCCTTCATGAGGAGGCTGAGTGGGGAGGAGGAGGAG 1286
QY 1093 GTGCTGAACTTCTCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1143
D 1094 GTGCTGAACTTCTCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1144
D 1095 GTGCTGAACTTCTCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1145
QY 1285 CCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1344
D 1286 CCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1345
D 1287 CCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1346
QY 1144 CGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1203
D 1145 CGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1204
D 1146 CGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1205
QY 1345 AGCTTCTAGGAAAGATATAGCTGCTTACCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1404
D 1346 AGCTTCTAGGAAAGATATAGCTGCTTACCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1405
D 1347 AGCTTCTAGGAAAGATATAGCTGCTTACCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1406
QY 1204 GGCCTGCACAGCACT---TATGATGCTTATTTCTGAGGAGGAGGAGGAGGAGGAGGAG 1260
D 1205 GGCCTGCACAGCACT---TATGATGCTTATTTCTGAGGAGGAGGAGGAGGAGGAGGAG 1261
D 1206 GGCCTGCACAGCACT---TATGATGCTTATTTCTGAGGAGGAGGAGGAGGAGGAGGAG 1262
QY 1405 CGGGTGTGCGATTTGAGGAGCTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1464
D 1406 CGGGTGTGCGATTTGAGGAGCTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1465
D 1407 CGGGTGTGCGATTTGAGGAGCTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1466
QY 1261 AAGAGAGTGCCTGAGCTCCAGATCCACATCATTTGAGGAGGAGGAGGAGGAGGAGGAG 1320
D 1262 AAGAGAGTGCCTGAGCTCCAGATCCACATCATTTGAGGAGGAGGAGGAGGAGGAGGAG 1321
D 1263 AAGAGAGTGCCTGAGCTCCAGATCCACATCATTTGAGGAGGAGGAGGAGGAGGAGGAG 1322
QY 1465 CCACAGCGGGTGTGAGGAGTGAATTTGACACAGATTTGCTGCTGCTGCTGCTGCTGCTG 1518
D 1466 CCACAGCGGGTGTGAGGAGTGAATTTGACACAGATTTGCTGCTGCTGCTGCTGCTGCTG 1519
D 1467 CCACAGCGGGTGTGAGGAGTGAATTTGACACAGATTTGCTGCTGCTGCTGCTGCTGCTG 1520
QY 1321 GGACAGCTGTGACAGACCTGCTTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
D 1322 GGACAGCTGTGACAGACCTGCTTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1381
D 1323 GGACAGCTGTGACAGACCTGCTTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1382
```


Db	693	CGAGCAGAGGATCGAGCTGGCACTCGTCAGGTGCCATCGGAGCTCTGCTACCTC	752
QY	363	CAAGACAGAGGAGGAATGTGCACAAATTTATCCAGATTCTGCCAATTGTCAATGCC	422
Db	753	GAAGTCAGAGGAGGAGCTGGCAGAACTACATCCAGTCTCTGGCGAAATTCACCAT	812

192	QY	GTCTCATACGTACAAATTACTCTGCTCTCCCTTGTGGATCCTGCCCTCCACACACACTTTACGT	251
193			
194			
195			
196	Db	GGCCACAGCTCCAGTTATCATCTCTTGGATGAGAACGGAGTAGGCTGTATGT	227
197			
198			
199			
200			
201	QY	CGGTGCACGGGATAGCATCTTCGGTTTAAACCTCCCTTCTCTGGGNAAGACCCGGAAG	311
202			
203			
204			
205	Db	TGGAGCAAAAGGATCACATATT--TTCTATCGACCTGGTTAATATCAAGGATTTTTCAAA	284
206			
207			
208			
209			
210	QY	GATCGACTGATGGTACCTGAGACTCACAGACACAACTGCAGGAGAAAGGCA--AGAA	368
211			
212			
213			
214	Db	GATTGTGTGCCAGTATCTTACACCAGGAGAGATGAATGGCAAGTGGGCTGGAAAAGACAT	344
215			
216			
217			
218	QY	AGAGGAGCAATGTACAAATTTATCCAGATTCTCGCCATTGTCAATGCTCTCACTCCT	428
219			
220			
221	Db	CTTGAAGAATGTCTAATTTCATCAAGGTACTTAAAGCATATATATCAGACTCACTTGTA	404
222			
223			
224	QY	CAGCTGGCGCACCTTCCTTTTGTATCCGAAGTGGGGTTATTGATGTCTCAGTTTCCA	488
225			
226			
227	Db	CGCTGTGGAACGGGGCTTTTCATCCAATTTGCACCTACATTTGAAATTTGACATCATCC	464
228			
229			
230	QY	GCAGGTTGAAGACTTGAG-----AGCGCGCGGGGAATGTCC	527
231			
232			
233	Db	TGAGGACAATATTTTAAAGCTGGGAGAACTCACATTTTCAAAACGGCGCTGGGAGAGTCC	524
234			
235			
236	QY	TTTTTGAGCCCAAGCGGTACAGGCTGAATGGCTGGGGCGTCTCTTACACGCCAC	587

Db 525 ATATGACCTTAAGCTCTGACAGCATCCCTTTTAATAGATGGAGAATTATATCTCTGGAAC 584
QY 588 TGTGAAGAACTTCTCTGGGAGCTGAGCCCATCATCTCCCGAGCTGTGGGTCGAGCTGAGGA 647
Db 585 TGCAGCTGATTTTATGGGGCGAGCTTTGCTATCTTCCGAACCTTTGGGCGACCAACCC 644
QY 648 CTGGATTGCAACAGACACCTTGTCTCTGGCTTAATGCTCCAGCCTTTGTGCGAGCTAT 707
Db 645 AATCAGGAGACAGCATGATTCAGGTGGCTCAATGATCCAAAGCTT-----CAAT 696
QY 708 GGTCTGAGCCAGCTGAGTGGGGGATGAAGATGGAGAGATGAATCTTTTTCCT 767
Db 697 AGTGCCCACTCATCTCAG-AGAGTGACAATCTCGAAGATGACAAAGTATATCTTTTCT 755
QY 768 CACGGAGACCTCCCGAGTGTGGAGCTCTATGAGCGCATCAAGTCCCAAGAGTGGCCCG 827
Db 756 CCGTGAATAATGCAATAGATGGAGACACTCTGGAAGCTACTCAGGCTAGATAGTCA 815
QY 828 AGTGTGTGGGGGACCTTGGGGGAGGAGACCTTTCAGCAGAGATGACGACGTTTCT 887
Db 816 GATATGCAAGAATGACTTTGGAGGGCAGAGAAGTCTGGTGAATAATGACAAACATCTCT 875
QY 888 GAAGGCTGACCTGTGTGCCAGGCGCCG 916
Db 876 CAAGCTCTGATTTGCTCAGTGCCAG 904

RESULT 13

US-09-060-692-53
; Sequence 53, Application US/09060692
; Patent No. 5935865
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,692
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET INFORMATION:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2601 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS
; LOCATION: 16..2331
US-09-060-692-53

Query Match 2.6%; Score 60.6; DB 2; Length 2601;

Best Local Similarity 47.9%; Pred. No. 1.3e-06;

Matches 359; Conservative 0; Mismatches 354; Indels 36; Gaps 5;

QY 192 GTCTCATAGTACAATTACTCTGCTCTCTGTCGATCTCCCTCCACACACATTTTACGT 251
Db 168 GGCACACAGCTCCAGTTATCATACCTTCTCTTTGGATGAGGAACGGAGTAGGTGTATGT 227
QY 252 CGGTGCACGGGATGAGCATCTTCGCTTTTAACCTCTCCCTCTCTCTGGGGAAGACCCCGAAG 311
Db 228 TGGAGCAAGGATCAGATATT---TTTCATTCGACCTTGGTAAATATCAAGGATTTTCAAAA 284
QY 312 GATCGACTGGATGGTACCTGAGACTCACAGACAGAACTCAGGAAGAAAGGCA---AGAA 368
Db 285 GATTGTGTGGCCAGTATCTTACACCAGAAGAGATGAATGCAAGTGGGCTGGAAGACAT 344
QY 369 AGAGGACGAATCTCAAAATTTTATCCAGATTCTCGCCATTTGTCAATGCCTCTCACCTCCT 428
Db 345 CCTGAAGAATGTGTAATTTTCATCAAGGTACTTAAGGCATATAATCAGACTCACTTGTA 404
QY 429 CACGTGCGGCACCTTCGCTTTTGTATCCGAAGTCCGGGGTATTGATGTGTCCAGTTTCCA 488
Db 405 CGCCTGTGAACGGGGCTTTTCATCCAATTTGCACCTACATTTGAATTTGGACATCATCC 464
QY 489 GCAGGTTGAAAGACTTTGAG-----AGCGGCGGGGGGAAATGTGCC 527
Db 465 TGAGGACAATATTTTAAAGCTGGAGAACTCACATTTTGAAGACGGCGCTGGGAAGAGTCC 524
QY 528 TTTTGAGCCAGCTCAACGCTCAGCAGCTGTAATGGCTGGGGGCTCTCTACACCCGCCAC 587
Db 525 ATATGACCTTAAGCTGCTCAGCAGATCCCTTTTAATAGATGGAGAATTATATCTCTGGAAC 584
QY 588 TGTGAAGAACTTCTCTGGGAGCTGAGCCCATCATCTCCGAGCTGTGGGTCGAGCTGAGGA 647
Db 585 TGCAGCTGATTTTATGGGGCGAGACTTTGCTATCTTCCGAACCTTTGGGCGACCAACCC 644
QY 648 CTGGATTGCAACAGACACCTTTGTCTCTGGCTTAATGCTCCAGCCTTTGTGCGAGCTAT 707
Db 645 AATCAGGACAGACAGCATGATTCAGGTGGCTCAATGATCCAAAGCTT-----CAAT 696
QY 708 GGTCTGAGCCAGCTGAGTGGGGGATGAAGATGGAGAGATGAATCTTTTTCCT 767
Db 697 AGTGCCCACTCATCTCAG-AGAGTGACAATCTCGAAGATGACAAAGTATATCTTTTCT 755
QY 768 CACGGAGACCTCCCGAGTGTGGAGCTCTTATGAGCGCATCAAGGTCCCAAGAGTGGCCCG 827
Db 756 CCGTGAATAATGCAATAGATGGAGACACTCTTGGAAAAGCTACTCAGGCTAGATAGTCA 815
QY 828 AGTGTGTGGGGGACCTTTGGGGGAGGAGACCTTTCAGCAGAGATGACGACGTTTCT 887
Db 816 GATATGCAAGAATGACTTTGGAGGGCAGAGAAGTCTGGTGAATAATGACAAACATCTCT 875
QY 888 GAAGGCTGACCTGTGTGCCAGGCGCCG 916
Db 876 CAAGCTCTGATTTGCTCAGTGCCAG 904

RESULT 14

US-08-833-391-53
; Sequence 53, Application US/08833391
; Patent No. 6013781
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family

NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,391
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/121,713
FILING DATE: 13-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 2601 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 16..2331
US-08-833-391-53

Query Match 2.6%; Score 60.6; DB 3; Length 2601;

Best Local Similarity 47.9%; Pred. No. 1.3e-06;
Matches 359; Conservative 0; Mismatches 354; Indels 36; Gaps 5;

QY 192 GTCTCATAGTACAAATTTACTGCTCTCTGTTGGTACCTGCTCCACACACTTTAGCT 251
DB 168 GCCAACAGCTCCAGCTTATCATACCTCTTTGGATGAGAACGGAGTAGGCTGTATGT 227
QY 252 CGGTCCAGGGATGATCATCTTGGCTTTAACCTCCCTCTCTCTGGGAAAGACCCGGAAG 311
DB 228 TGGAGCAAGATCATCATATT---TTCATTGACCTGGTGAATCAAGGATTTTCAAAA 284
QY 312 GATCGACTGGATGGTACCTGAGACTCAGACAGACTGCAGGAAGAAGCA---AGAA 368
DB 285 GATTGTGGCCAGTATCTTACACAGAGAGATGAATGCAAGTGGGCTGGAAAAGACAT 344
QY 369 AGAGCAAGATGTCAAAATTTATCCAGATTTCTGCCATTTGTCATGCTCTCACCTCCT 428
DB 345 CCTGAAGAGATGTGCTAATTTATCAAGGACTTAAGGCATATATCATGACTCATTGTA 404
QY 429 CAGGTGGCGGACCTTTCGCTTTTGATGCCAAGTGGGGGTTATTGATGTGTCCAGTTTCCA 488
DB 405 CGCCTGTGGAACGGGGGCTTTTCATCCAAATTTGCACCTTACATTGAAATTTGGACATCATCC 464
QY 489 GCAGTTTGAAGACTTGAG-----ACGCGCGGGGGAATGTCC 527
DB 465 TGAGGACAATATTTTAAGCTGGGAGAACTCAATTTTGAACACGGCGCTGGGAGAGTCC 524
QY 528 TTTTGAGCAGCTCAACGGTCAGCAGCTGTATGGCTGGGGGCTGCTCTACACCGCCAC 587
DB 525 ATATGACCTTAACTGCTGACAGCATCCCTTTTATAGATGAGAAATATATCTTGGAAC 584
QY 588 TGTGAAGAACTTCCTTGGGGAGTGAGCCCATCATCTCCCGAGCTGTGGTGCAGCTGAGGA 647

DB 585 TGCAGCTGATTTTATGGGGCGAGACTTTGCTATCTTCGAACTCTTGGCACCACCC 644
QY 648 CTGGATTCCAGACAGACCTTTGCTATCTGCTTAACTGCTCCAGCCTTTGTCGACGCTAT 707
DB 645 AATCAGGACAGACGACATGATTCAGGTGCTCAATGATCCAAAGTT-----CAT 696
QY 708 GGTCTGAGCCAGCTGAGTGGGGGATGAAGATGGACGATGAAATCTTTTCTTCTT 767
DB 697 AGTGGCCACCTCATCTCAG-AGAGTGACAATCCTTGAAGATGACAAGTATATCTTTTCTT 755
QY 768 CAGGAGACCTCCCGAGTGTGGACTCTATGAGCGCATCAAGTCCCAAGAGTGGCCG 827
DB 756 CCGTGAATAATGCAATAGATGGAGAACACTCTGGAAGAGCTACTCAGCTAGAAATAGTCA 815
QY 828 AGTGTGTGGGGGACCTTGGGGGAGGAAGACCTTCAGCAGAGATGGACGACCTTCT 887
DB 816 GATATGCAAGATGACTTTTGGAGGACACAGAGTCTGGTGAATAAATGACAACATCTCT 875
QY 888 GAAGCTGACCTGCTGTCGCCAGGCGCG 916
DB 876 CAAAGCTGCTGATTTGCTCAGTCCAG 904

RESULT 15

US-09-060-610-53
Sequence 53, Application US/09060610
Patent No. 634544
GENERAL INFORMATION:
APPLICANT: Goodman, Corey S.
APPLICANT: Kolodkin, Alex L.
APPLICANT: Matthes, David
APPLICANT: Bentley, David R.
APPLICANT: O'Connor, Timothy
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/060,610
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/835,268
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 2601 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 16..2331

Query Match	2.6%	Score 60.6;	DB 4;	Length 2601;
Best Local Similarity	47.9%	Pred. No. 1.3e-06;		
Matches 359;	Conservative 0;	Mismatches 354;	Indels 36;	Gaps 5;
QY	192	GTCTCATACGTACAAATTACTCTGCTCTCCCTTGTGGATCCTCGCTCCACACACTTTTACGT	251	
Db	168	GGCCACAGCTCCAGTTATATCATCTTCTTTGGATGAGAACGGAGTAGGCTGTATGT	227	
QY	252	CGGTGCACGGATAGCATCTTCGCTTTTAAACCCCTCCCTTCTCTGGGAAAGACCCCGAAG	311	
Db	228	TGGAGCAAAAGATACATATAFT---TTCATCGACCTGGTTAATATCAAGGATTTTCAAA	284	
QY	312	GATCGACTGGATGGTACCTAGACTCACACAGAACTGCAGGAAGAAAGGCA---AGAA	368	
Db	285	GATTTGTGTGCCAGTATCTTTACACCAAGAGAGATGAATGCAAGTGGCTGGAAAAGACAT	344	
QY	369	AGAGGACGAATGTACAAATTTTATCCAGATTTCTGCCATTGTCAATGCCCTCTCACCTCCCT	428	
Db	345	CCGTGAAGATGTGCTTAATTTATCAAGGTACTTTAGGCGATATAATCAGACTCACTTGTGA	404	
QY	429	CAGCTGGGCACCTTCGCTTTTGATCGGAAGTCGGGGTTTATTGATGTGCCAGTTTCCA	488	
Db	405	CGCTCTGGAAACGGGGCTTTTCATCCAAATTTGCACCTACATTTGAAATTTGGACATCATCC	464	
QY	489	GCAGGTGTAAGACTTTGAG-----AGCGCGCGGGGAAATGTCTCC	527	
Db	465	TGAGGACATATTTTTTAAGCTGGAGAACTCACATTTTGAAGACGGCGTGGGAAGAGTCC	524	
QY	528	TTTTTGACCCAGCTCAACGGTTCAGAGCTGTAAATGGCTGGGGGGCTCTCTACACCGCCAC	587	
Db	525	ATATGACCCTAAGCTGCTGACAGCATCCCTTTTAAATAGATGGAAATTAATCTCTGGAAC	584	
QY	588	TGTGAGAACTTCTCGGGNACTGAGCCCATCATCTCCGAGCTGTGGTGCAGCTGAGGA	647	
Db	585	TGCAGCTGATTTATGGCGGAGACTTTGCTATCTTCCGAACCTTTGGGCACCAACCC	644	
QY	648	CTGGATTCCAACAGAGACCTTGTTCATCTCGGCTTAAATGCTCCAGCTTTTGTCCGACGTAT	707	
Db	645	AATCAGACAGACGACGATGATTTCCAGGTGGCTCAATGATCCAAAGTT-----CATT	696	
QY	708	GGTCTTGAGCCAGCTGAGTGGGGGATGAAGATGGAGACGATGAAATCTTTTTTTTCTT	767	
Db	697	AGTGCCCACTCATCTCAG-AGAGTGACAACTCTGAAGATGACAAAGTATACTTTTTCTT	755	
QY	768	CACGGAGACTCCCGAGTGTGTGACTTCTATGAGCGCATCAAGTCCCACAGATGGCCCC	827	
Db	756	CCGTGAAATGCAATAGATGGAGAACACTCTCTGAAAAGCTACTTCAGCGTAGAATAGGTCA	815	
QY	828	AGTTGTGCGGGGACCTTTTGGGGCAGGAAGACCCCTTCAGCAGAGATGGACGACGTTTCT	887	
Db	816	GATATCAAGAATGACTTTTGGAGGCACAGAAGTCTTGGAAGCTACTTCAGCGTAGAATAGGTCA		
QY	888	GAAGGCTGACCTGTGCGCCAGGGCCCG	916	
Db	876	CAAAAGCTCGTCTGATTTGCTCACTGCGAG	904	

Search completed: September 26, 2003, 05:46:34
Job time : 128.291 secs ,

SUMMARIES									
Result No.	Query ⁸			Description					
	Score	Match	Length	DB	ID				
1	1626	69.8	2893	11	US-09-808-665A-1				Sequence 1, Appli
2	1626	69.8	2894	11	US-09-808-665A-3				Sequence 3, Appli
3	654	28.1	799	11	US-09-808-665A-5				Sequence 5, Appli
4	567.6	24.4	928	14	US-10-188-246-13				Sequence 13, Appl
5	527.4	22.6	777	14	US-10-188-246-15				Sequence 15, Appl
6	251.8	10.8	3293	14	US-10-149-819-25				Sequence 25, Appl
7	248.6	10.7	2155	13	US-10-003-152-11				Sequence 11, Appl
8	248.6	10.7	2155	14	US-10-002-050-11				Sequence 11, Appl
9	248.6	10.7	2155	14	US-10-002-304-11				Sequence 11, Appl
10	248.6	10.7	2156	13	US-10-003-152-21				Sequence 21, Appl
11	248.6	10.7	2156	14	US-10-002-050-21				Sequence 21, Appl
12	248.6	10.7	2156	14	US-10-002-304-21				Sequence 21, Appl
13	248.6	10.7	2284	13	US-10-003-152-13				Sequence 13, Appl
14	248.6	10.7	2284	14	US-10-002-050-13				Sequence 13, Appl
15	248.6	10.7	2284	14	US-10-002-304-13				Sequence 13, Appl
16	247	10.6	3781	11	US-09-946-374-252				Sequence 252, App

QY	121	GTCCCCCGCTCAGTGGCCAGAACCTGGCTGCCAATCTCCGAGGCTGACTCTATCTCAACC	180
DB	215	GTCCCCCGCTCGGTGCCAGAACCTCGCTTCCAATCTCTGAGGCTGACTCTGTCTCAACC	274
QY	181	CGGTTTGAGCGTCTCATACCTACAAATTACATCTCTCTCCPTGTGGATCCTGCCTCCAC	240
DB	275	CGGTTGCGAGTCCCTCACACATACAAATTACTCTGTCTCTCTGTGGATCTGCTCTCCAC	334
QY	241	ACACTTTACGTCGGTGCACGGATAGCATCTTTCGCTTTAAACCTTCCCTCTCTCTGGGAA	300
DB	335	ACACTTTATGTTGGCCCGGACACCATCTTCGCTTTATCCCTTGCCCTTCTCAGGGGAG	394
QY	301	AGACCCGAAGAGTCGACTTGATGGTACCTGAGACTCACAGACAGAACTGCAGGAAGAAA	360
DB	395	AGACCCCGCAGGATTTGACTGATGGTTCTCTGAGGCTCACAGACAGAACTGTAGGAAGAAA	454
QY	361	GGCAAGAAAGAGGACGAATGTCACAATTTATCCAGATTCTCGCCATTGTCAAATGCCTCT	420
DB	455	GGCAAGAAAGAG-----	466
QY	421	CACCTCCTACGTCGCGCACCTTCGCTTTTGATCCGAAGTCGGGGTTAATGATGTGTCC	480
DB	467	-----GATGTGTCC	475
QY	481	AGTTTCCAGCAGGTTGAAGACTTGCAGCGGCCGGGGAAATCTCCTTTTGAGCCAGCT	540
DB	476	AGGTTCCAGCAGGTTGAAGACTTGCAGATGGCCGGGGAAATCTCCTTTTGAGCCAGCT	535
QY	541	CAACGGTCAAGCAGCTGTAATGGCTGGGGCTCTCTACACCGCCACTGTGAAGAAGCTTC	600
DB	536	CAGCGGTCAAGCAGCTGTAATGGCTGGGGGGTCTCTATGCTGCCATCTGTGAAGAACTAC	595
QY	601	CTGGGGACTGAGCCCATCATCTCCGAGCTGTGGGTCGAGCTGAGGACTGGATTCGAACA	660
DB	596	CTGGGACGGAGCCAATTTATCACAGAGCAGTGGGTCGTGCCGAGGACTGGATTCGGACA	655
QY	661	GAGACCTTGTCATCCTGGCTTAATGCTCCAGCCCTTTGTGGAGCTATGCTCTGAGCCCA	720
DB	656	GATACCTTGCCTCCTGGCTGAACGCCCCAGCCCTTTCGACGCGGTGGGCTTGAGCCCA	715
QY	721	GCTGAGTGGGGGATGAAGATGGAGACGATGAATCTTTTCTTTCACGGAGACCTCC	780
DB	716	GCCGAATGGGGGATGAAGATGGAGACGACGAATCTACTTCTTTACGGAGACTTCC	775
QY	781	CGAGTGTGGACTCCTATGAGCGCATCAAGGTCCCAAGATGGCCCCGAGTGTGTGCGGG	840
DB	776	CGAGCTTTGACTCATACGAGCGCATTTAAAGTCCACGGGTGGCCCCGTGTGTGCGGGG	835
QY	841	GACCTTGGGGCAGGAAGCCCTCACAGAGATGGACGAGTTTCTGAAGGCTGACCTG	900
DB	836	GACCTCGGGGGCCGGAAGACCTCCACGAGATGGACGAGTTTGAAGCTGACCTG	895
QY	901	CTGTGCCAGGGCCGAGCATGGCCGGCCCTCCGGGGTTCTGCAGGCTATGGCAGAGCTT	960
DB	896	CTCTGTCCAGGGCTGAGCATGGCCGGCCCTCCAGTGTCTCTGCAGGATGTGTGTGCTT	955
QY	961	CGGCTCAGCCTTGAGCGGGAAACCCCATCTTTTATGGGATCTTTTCTCCCAAGTGGAA	1020
DB	956	CGACTGAGCTTGGGGCAGGGAGTCCCATCTTTTATGGCATCTTTTCTCCCAAGTGGAG	1015
QY	1021	GGAGCTGCCATCTCTGCTGTGTGCTTCCGACCCCAAGACATCCGGGAGTGTCTCAAT	1080
DB	1016	GGGCTACTTCTCTGCTGTGTGCTTCCGACCCCAAGACATCCGGAGTGTCTGAAAT	1075
QY	1081	GGTCCCTTTAGAGAGCTTAAACATGACTGCAACAGGGGAGTGCCTGTCTATGGACACGAG	1140
DB	1076	GGTCCCTTCAGAGAAGTAAACATGACTGCAACAGAGGACTGCCTGTCTGTGGACATGAT	1135
QY	1141	GTGCCCAGCCAGACCTTGAGAGTGTGATCGCCAAACATGAAGTCTCAGCAGTTTGA	1200
DB	1136	GTGCCCAGCCAGACCTTGAGAGTGTGATCACCACAAACATGAAGTCTCAGCAGTTTGGC	1195

QY	1201	TCCTCACTCCTCCCTGCCAGACCGCGTGCTCACTTATTAACAGAACACCCTCTCATGGAC	1260
Db	1196	TCATCTCTCTCCCTGACCGCGTACTCACCTTATCCCGGACACCCACTCATGGAC	1255
QY	1261	AGGCCCGTTTCCCAGGTGAGCGCGCCCGCTGCTGGTGACTACAGATACAGCCTATCTC	1320
Db	1256	AGGCCAGTGTTCACAGTGAATGGGCACCCCGCTGCTGGTGACTACAGATACAGCCTATCTC	1315
QY	1321	AGAGTCGTGGCCCCACAGGCTGACCAGCGCTCTCAGGGAAAGAATAATACAGTGCCTACCTG	1380
Db	1316	AGAGTCGTGGCCCCACAGGCTGACCAGCGCTCTCAGGGAAAGAATAATACAGTGCCTACCTG	1375
QY	1381	GGGACAGAGGATGGACACCTCCACCGGGCTGTGCGATTGGAGCTCAGCTAGTGTCTTTG	1440
Db	1376	GGGACAGAGGATGGACACCTCCACCGAGCAGTGGGATCGGAGCTCAGCTCAGCGTCTCTT	1435
QY	1441	GAGGATCTGGCCTTGTTCACAGAACACACGCGGGTTGAGAGCATGAANTGTACACAGAT	1500
Db	1436	GAGGATCTGGCCTTGTTCACAGAACACACGCGGATGGAGCAATGAANTGTACACAGC	1495
QY	1501	TGGCTCTCTGTGGGCTCCCACTACTGAGGTGACACAAGTGAACACCACTCTGGCCGT	1560
Db	1496	TGGCTCTCTGTGGCTCCCGTACTGAGGTGACACAAGTGAANTGTACACCACTCTGGCCGT	1555
QY	1561	CTCAGAGCTGCTCGGAGTGTATCTCTGGCCCCAGAACCCGCTGTGCGCTGGAGTTCGCG	1620
Db	1556	CTCAGAGCTGCTCAGAGTGCATCTCTGGCCCCAGAACCCAGTCTGTGCTGGAGTTCGCG	1615
QY	1621	CTTGTATGCTTGTGGCCCCAGCGCGGAGACCGCGGGATGTTCAAAGATATAGAGTCA	1680
Db	1616	CTGTATGAGTGTGGCCCCATGCGGGGAGCACCGAGGGTTGGTCCAAGACATAGAGTCA	1675
QY	1681	GCGGATGCTCTCTTCTTGTGTCCAAAAGAACCTGGAGAACATCCCGTAGTGTGTAAGTT	1740
Db	1676	GCAGATGTCCTCTTGTGTCTTAAGAGCCTGGAGAGCTCCAGTAGTGTGTAAGTT	1735
QY	1741	CCGCTGCTACTGTGGCCCCAGCTGGTGCTGCCATGTTCGCCCAAGTCTGCTCGCTGGGCATCC	1800
Db	1736	CCCCTGGCTACAGCTGCGCATGTGGTCTTGCCATGTTCCTCAAGCTCAGCATGGGCATCC	1795
QY	1801	TGTTGTGGCACACAGCCCCAGTGAGTGACTGCGCTCAGCTCCCGGAGGGATGGACTAGAG	1860
Db	1796	TGTTGTGGCACACAGCCCCAGTGAGTGACTGCACTCACCCCCCGCGGATGGACTGAG	1855
QY	1861	GTGTTGTGACCCACAGGGCCATGGGGCTTATGCTTTGCGAGTGTCAAGGAGGTGGAGCC	1920
Db	1856	GTGTTGTGACCCACAGGGCCATGGGGCTTATGCTTTGCGAGTGTCAAGGAGGTGGGCA	1915
QY	1921	GCCCGCTGTGGCTGTATAGCTTTGGTGTGGGGAGCCAGCGGGACCCCTCAACCGG	1980
Db	1916	GCCCATGTGTTAGCAGCTTACAGTTGGTATGGGGCAGCCAGCGAGATGCTCCGAGCCGG	1975
QY	1981	GCCACACCGTTGTGGGGCTGGATTTGGTTGGCTTTCTCTCTGGGTGTTCTTCAGCATCC	2040
Db	1976	GCCACAC --- AGTGGGGCGGAGCTGGCTGGGTCTCTTGGGGATTTCTCGACGATCC	2032
QY	2041	CTCACTCTCTCTGATTGTGTCGCGCTACAGCGCTCGGCGACAGAGGAGCTTCTAGCT	2100
Db	2033	CTGACTCTCATTTGATTGTTCCGGCTCAGCAGCGGAGCGGACAGAGGAGACTCTTGGT	2092
QY	2101	AGACAGAGGTGGGCTTATAGATCTGGGGGCTCCACCTTCTGGGACCAACAGCTATAGTCAG	2160
Db	2093	AGACAGAGGTGGGCTTGGACCTGGGGGCTCCACCTTCTGGGACCAACAGCTACAGCCAA	2152
QY	2161	GACCTTCCCTCTCTTGGCTGAAGATGAACGCTGCCCTGGCCCTGGGTAAAGGGGG	2220
Db	2153	GACCTTCCCTCTCTTGGCTGAAGATGAACGCTGGCCCTGGCCCTGGCCAAAGAGGGG	2212
QY	2221	AGTGGTTTTGGTGGCTTCCCTCCACCTTCTCTCTGGATTCTTGGCCAAAGCCAGCCAC	2280
Db	2213	AGTGGTTTTGGTGGATTCTCACCACTTCTCTCTGCTGTGATCTTGGCCAAAGCCAGCCAC	2272
QY	2281	ATCCGGTCACTGGGGCGCTCTTAGCCAGGTGTGATGAGACCTTCCATCTA2330	

1616	CTGGATGAGTCTGTGGCCCATCGCGGGGACCGAGGGTTGTCTCAAGACATAGAGTCA	1675
1681	GCGGATGTCCTCTCTTTGTGTCCAAAGAACCCTGGAGAACATCCCGTAGTGTGTTGAAGTT	1740
1676	GCAGATGTCCTCCCTCTTTGTCTCTAAAGAGCCCTGGAGAACGTCCTCAGTAGTGTGTTGAAGTT	1735
1741	CCGGTGGCTACTGTGGGCCACGTGGTCTCTGCCATGTTCCCCAGTTCTGCCTGGGCATCC	1800
1736	CCCGTGGCTACAGCTGGCATGTGGTCTTGCCCATGTTCTCAAGCTACAGCATGGGCATCC	1795
1801	TGTGTGTGGCACCAGCCCACTGGAGTGACTGCGCTCACTCCCGGAGGGATGACCTAGAG	1860
1796	TGTGTGTGGCACCAGCCCACTGGAGTGACTGCACTCACCCCCCGCGGGATGACCTGGAG	1855
1861	GTGGTGGTGACCCAGGGGCCATGGGGGCTTATGCTTTCGAGTGTTCAGGAGGTGGAGCC	1920
1856	GTGGTGGTGACCCAGGGGCCATGGGCGCTTATGCTTGAATCTCAGGAGGTGGGGCA	1915
1921	GCCCGCGTGTGCTGCTTATAGCTTGTGTGGGCGAGCCAGCGGGGACCCTCAAAACCGG	1980
1916	GCCCATGTGTACGACTTACAGCTTGGTATGGGGCAGCCAGCGAGATGCTCCGAGCCGG	1975
1981	GCCCAACACCGTGTGGGGGTGGATTGGTGGCTTCTCTCTGGGTCTTCTTTCGAGCATCC	2040
1976	GCCCAACAC---AGTGGGGGGGACTGGCTGGCTTCTTCTTGGGGATCTTCGCAGCATCC	2032
2041	CTCACTCTCTCTGTGTCGGCTGCGCTCAGCAGCGTCGGCGCAGAGGGAGCTTCTAGCT	2100
2033	CTGACTCTCAATCTGAATGTGTGGCGTTCAGCGACGCGGGCGACAGAGGAACTTCTGGCT	2092
2101	AGAGCAAGGTGGGCTTAGACTCTGGGGGCTCCACCTTCTGGGACCACAAGCTATAGTCAG	2160
2093	AGAGCAAGGTGGGCTGGACCTGGGGCTCCACCTTCTGGGACCACAAGCTACAGCCAA	2152
2161	GACCTCCCTCTCCTTCGGCTGAAGATGAACGCGTGCCTCGCCCTGGGTAGCGGGGC	2220
2153	GACCTCCCTCCCTCTCTCTGAAGATGAGCGGTTGCCGTGGCCCTGGGCAAGAGGGGC	2212
2221	AGTGGTTTTGGTGGCTTCCCTCCACCTTCTGCTGGATTCCTTTCGCCAAGCCAGGCCAC	2280
2213	AGTGGCTTTGGTGGATTCACACACCTTCTCTGCTTGATCTTTCGCCAAGCCAGGCCAC	2272
2281	ATCCGGCTCACTGGGGGCGCTCTAGCCACGTGTGATGAGACCTTCCACTTA	2330
2273	ATTCCGGCTAACTGGGGCTCTCTTAGCCACATGTGATGAACATCCACTTA	2322

RESULT 3

```

US-09-808-665A-5
; Sequence 5, Application US/09808665A
; Publication No. US2003002282A1
; GENERAL INFORMATION:
; APPLICANT: David Michalovich
; APPLICANT: Trudy Rachel Doe
; APPLICANT: Philip David Hayes
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP-30036-C1
; CURRENT APPLICATION NUMBER: US/09/808,665A
; CURRENT FILING DATE: 2001-03-15
; PRIORITY APPLICATION NUMBER: 09/160,762
; PRIORITY FILING DATE: 1998-09-24
; PRIORITY APPLICATION NUMBER: 98300693.3
; PRIORITY FILING DATE: 1998-01-30
; PRIORITY APPLICATION NUMBER: 9816423.9
; PRIORITY FILING DATE: 1998-07-28
; PRIORITY APPLICATION NUMBER: 9816676.2
; PRIORITY FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 799
; TYPE: DNA
; ORGANISM: HOMO SAPIENS

```

; FEATURE:									
; NAME/KEY: UNSURE									
; LOCATION: (720)									
; OTHER INFORMATION: OTHER INFORMATION: Partial EST derived sequence									
US-09-808-665A-5									
Query Match									
Best Local Similarity 28.18; Score 654; DB 11; Length 799;									
Matches 708; Conservative 0; Mismatches 91; Indels 0; Gaps 0;									
Qy	1051	CGACCCCAAGACATCCGGGCAGTCTGCTGAATGGTCCCTTTAGAGAGCTAAACATGACTGC	1110						
Db	1	CGACCACAGACATTCGGACAGTCTGCTGAATGGTCCCTTCAGAGACTAAACATGACTGC	60						
Qy	1111	AACAGGGAGTGGCTGTCTATGGACAAACGAGGTGCCAGCGCCACAGACTGGAGAGTGCATC	1170						
Db	61	AACAGAGGACTGCCTGTGCTGGACAAATGATGTGCCAGGCCACAGACTGGAGAGTGCATC	120						
Qy	1171	GCCAAACAACATGAGCTCCAGCAGATTGGATCCTCACTCTCCCTGCCAGACCGCGTCTC	1230						
Db	121	ACCAACAACATGAAGCTCCGGCACTTTGGCTCATCTCTCCCTGCCCTACCCGCTACTC	180						
Qy	1231	ACCTTTATCAGAGACCACCTCTCATGGACAGGCCCGCTGTTCCTGGCTGACGCCGCC	1290						
Db	181	ACCTTCATCGGGACCACCACTCATGGACAGGCCAGTGTTCAGCTGATGCCACCCC	240						
Qy	1291	CTGCTGTCTACTACAGATACAGCCTATCTCAGAGTCTGTGCCACACAGGTGACACAGCTC	1350						
Db	241	CTGCTGTCTACTACAGTTACAGCCTATCTCAGAGTCTGTGCCACACAGGTGACACGCTC	300						
Qy	1351	TCAGGGAAGAAATATGACGTGCTCTACCTGGGGACAGAGATGGACACTTCCACCGGGCT	1410						
Db	301	TCAGGGAAGAGATATGATGTGCTCTACCTGGGGACAGAGATGGACACTTCCACCGAGCA	360						
Qy	1411	GTGGCAATTGGAGCTCAGCTCAGTGTCTTGAGAGATCTGGCCTTGTTCGCCAGAACACAG	1470						
Db	361	GTGGGATCGGAGCTCAGCTCAGCTTCTTGAAGATCTGGCCTTATTCGCCAGGCCACAG	420						
Qy	1471	CCGCTTGAGAGCATGAAATTTGTACCAAGATTGGCTCTCTGTGGGCTCCCAATCTAGAGTG	1530						
Db	421	CCAGTTTGAGAACATGAAATTTGTACCAAGCTGGCTCTCTGGTTGGCTCCCTACTGAGGTG	480						
Qy	1531	ACAAAGTGAACACCAGCACTGTGGCCGTCCTCAGAGCTGCTCGAGTCTATCTCTGGCC	1590						
Db	481	ACAAAGTGAATACAACTGTGGCCGTCCTCAGAGCTGCTCAGAGTGCATCTCTGGCC	540						
Qy	1591	CAGGACCCCGTGTGCGCTGGAGCTTCGGCTTCGATCTTGTGTGGCCCAACGCCGCGAG	1650						
Db	541	CAGGACCCAGTCTGTGCCCTGGAGCTTCGGCTTCGATCTTGTGTGGCCCAATGCCGGAG	600						
Qy	1651	CACGCGGGATGGTTCAAGATATAGAGTACAGCGATGCTCTTCTTTGTGTCCAAAGAA	1710						
Db	601	CACGCGGGTTGGTCCAAAGACATAGAGTACAGAGATGCTCTCTTTGTGTCTCTAGAGAG	660						
Qy	1711	CTGTGGACAACATCCCGTAGTGTTTGAGTTCCCGTGGCTACTGTGGGCCAGGTGGTCTTG	1770						
Db	661	CTGTGGAGAAGCTCCAGTAGTGTTTGAGTTCCCGTGGCTACAGCTGCGCATGTGGTCTTN	720						
Qy	1771	CCATGTTTCCCGCAGTTCCTGCTGGGCATCTCTGTGTGGCACCAGCCAGTGGAGTGAAT	1830						
Db	721	CCATGTTTCCAGCTCAGCATGGGCATCTCTGTGTGGCACCAGCCAGTGGAGTGAAT	780						
Qy	1831	CGCGTCACTCCCGGAGGG	1849						
Db	781	TCACCTTACCCCGCGGG	799						

RESULTS

US-10-188-246-13
; Sequence 13, Application US/10188246
; Publication NO. US2003008274A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.,

[illegible]

Db 1490 GCACAGAGAGCGGTGGCTGCTCAAGGCTGTGAGCCTGGGGCCCTGGGTTCACCTGATG 1549
Qy 1442 AGGATGTGGCTTGTGTT---CCCAAGAACACAGCCGGTTGAGAGCATGAAATTTAGCACG 1498
Db 1550 AGGAGCTGCAGCTGTTTGACCAAGAGCCCATGAGAGCCCTGGTGTCTATCTCAGAGCAAGA 1609
Qy 1499 ATTGGCTCCTGGTGGGCTCCCATCTAGAGGTGACACAAGTGAACACCAAGCAACTGTGGCC 1558
Db 1610 AGCTGCTCTTTGGCCGCTCCCGCTCTCAGCTGTGAGCTGCGCGTGGCGGACTGCATAA 1669
Qy 1559 GTCTCAGAGCTGCTGGAGTGTATCTTGGCCAGGACCCCGTGTGGGCTTGAGGCTTCC 1618
Db 1670 AGTATCGCTCTGTGACAGATGTGCTCGCCGGGACCCCTATTGGCCTTGGAGGTCA 1729
Qy 1619 GGCTTGATGTTGTGGGCCACAGCCGGCGA 1649
Db 1730 ACACAGCCGCTGTGTGGCGTGGGTGGCCA 1760

RESULT 8

US-10-002-050-11
: Sequence 11, Application US/10002050
: Publication No. US20030032095A1
: GENERAL INFORMATION:
: APPLICANT: Shimkets, Richard
: APPLICANT: Fernandes, Elma
: APPLICANT: Vernet, Corine
: APPLICANT: Yang, Meijia
: APPLICANT: Boldog, Ferenc
: APPLICANT: Herrmann, John
: TITLE OF INVENTION: No. US20030032095A1el Nucleic Acid Sequences Encoding Human Semaph
: FILE REFERENCE: 15966-534 Cura-54 CON-S14
: CURRENT APPLICATION NUMBER: US/10/002,050
: CURRENT FILING DATE: 2001-11-02
: PRIOR APPLICATION NUMBER: 09/604,286
: PRIOR FILING DATE: 2000-06-22
: PRIOR APPLICATION NUMBER: 60/140,584
: PRIOR FILING DATE: 1999-06-23
: NUMBER OF SEQ ID NOS: 49
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 11
: LENGTH: 2155
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (166)..(1935)
US-10-002-050-11

Query Match 10.7%; Score 248.6; DB 14; Length 2155;
Best Local Similarity 52.1%; Pred. No. 8e-63;
Matches 704; Conservative 0; Mismatches 629; Indels 18; Gaps 6;
Qy 308 GAAGGATGCAGTGGATGTTACCTGAGACTCAGACAGAGACTGCAGGAGAAGCAAG- 366
Db 419 GAGCGATCTCTGGGAGGCCCGCTGGAGAAGAGACTGAGTATCCAGAAAGGGAAGA 478
Qy 367 --AAAGAGACGAATGTCAATAATTTATCCAGATTCCTGCCATTTGTCAATGCTCTCAC 424
Db 479 ACAACAGACCGAGTCTTCAACTTCATCCGTTCTGTGAGCCCTACAATGCTCCACC 538
Qy 425 TCCTCAGGTGGGACCTTCGCTTTTGTATCCGAAAGTGGGGGTTAATGATGTCCAGTT 484
Db 539 TGTACGTCGTGACACTACGGCTTCCAGCCCAAGTGCACTACGTCAACATGCTCACCT 598
Qy 485 TCCA---GCAGGTGGAAGACTTGAGAGCGCCGGGGAATGTCTTTTGAGCCAGCTC 541
Db 599 TCATTTGAGCATGGAGAGTTTGAAGATGGGAAGGCAAGTGTCCCTATGACCCAGCTA 658
Qy 542 AACGGTCAGCAGCTGTAATGGCTGGGGGGTCTCTACACCCGCACTGTGAAGAAGTCC 601
Db 659 AGGGCCATGCTGGCCTTCTGTGGATGGTGTACTGTACTGCGCCACACTCAACAAGTCC 718

Qy 602 TGGGGACTGAGCCCATCATCTCCCGAGCTGTGGGTGAGCTGAGACTGGAATTCGAACAG 661
Db 719 TGGGCACGGAACCCATTATCTTCGTAACATGGGGC---CCCACCACTCCATGAACACAG 775
Qy 662 AGACCTTGTCATCTCGCTTAATGCTTCCAGCCCTTTGTGGAGCATATGTTCTGTAGCCCCAG 721
Db 776 AGTACCTGGCCTTTTGGCTCAAGCAACCTCACTTTGTAGGCTCTGCTATGTAACCTGAGA 835
Qy 722 CTGAGTGGGGGATGAAGATGAGACGATGAATCTTTTTTTTCTTTCAGGAGACCTCC 781
Db 836 GTGTGGGAGCTTTCAGGGGGAGGACGACAAAGTCTACTTCTTCTCAGGAGCGGGCAG 895
Qy 782 GAGTGTGGAGTCTTATGAGCGCATCAAGGTCCCAAGAGTGGGCCGAGTGTGTGCGGGG 841
Db 896 TGGAGTCCGACTGCTATGCGGAGCAGGTGTGTGCTGCTGTGCGCCGCTGCTGCAAGGGG 955
Qy 842 ACCTTGGGGGAGGAAGACCTTTCAGCAGAGATGAGAGAGGTTCCTGAAGGCTGACCTGC 901
Db 956 ATATGGGGGGCGACGAGCCCTGCAGAGGAAGTGGACCACTTCTTCTGAAGCGCGCTGG 1015
Qy 902 TGTGCCACGGGCCGAGCATGGCCGGCTTCCGGGTTCCTGCAAGGCTATGGCAGAGCTTC 961
Db 1016 CATGCTCTGCCCGAAGTGGCAGCTCTACTTCAACAGCTGCGAGGCGATG---CACACC 1072
Qy 962 GGCCTCAGCTGGAGCGGAACCCCATCTTTATGGGATCTTTTCTCCAGTGGGAAG 1021
Db 1073 TGCAGGACACCTCTCTGGCACAACACCACTCTTCTTGGGGTTTTCAGACACAGTGGGTG 1132
Qy 1022 GAGCTGCCATCTCTGCTGTGTGCTTCCGACCCCAAGACATCCCGGCGAGTCTCAATG 1081
Db 1133 ACATGTACCTGTGGCCCATCTGTGTAGTACCAGTGTGAAGAGATCCAGCGGGTGTGAGG 1192
Qy 1082 GTCCCTTTAGAGAGCTAAAAACATGACTGCAACAGGGGACTGCTGTCTCATGGCAACGAGG 1141
Db 1193 GCCCTTATAAGAGTACCATGAGGAGCCGAGAGTGGGACCGCTACACTGACCTT--G 1249
Qy 1142 TGCCCGAGCCAGACCTGGAGAGTGCATCGCCAACAACATGAAGTCCAGAGTGTGGAT 1201
Db 1250 TACCCAGCCCTCGGCTGTGTCATTTAAACAACGTGGCATGGGCCACGCTACACCA 1309
Qy 1202 CTTCACTCTCCCTGCCAGACCGGCTGCTCACCTTTATCAGAGACCACTCTCATGGACA 1261
Db 1310 GCTCCCTGAGCTACCCGACAACTCTCAACTTCGTCAAGAAGCACCCGCTGATGGAG 1369
Qy 1262 GCGCGCTGTCGCGGTGACGCGCCCTCTGCTGTCTACTACAGATACAGCTATCTCA 1321
Db 1370 AGCAGTGGGGCTCGGTGAGCGCCCTCTGCTGTGAAGAGGCAACCACTTCACCC 1429
Qy 1322 GAGTGTGGCCACAGGTTGACAGCTCTCAGGGAAGAAATATGAGTGTGCTTACTGG 1381
Db 1430 ACCTGTGGCCGACCGGTTTACAGGACTTGTATGGAGCCACCTATACAGTGTCTCATTTG 1489
Qy 1382 GCACAGAGATGGACACCTCCACCGGGCTGTGGCATTTGGAGCTCAGTCTGCTTGG 1441
Db 1490 GCACAGGAGACGGCTGGCTGCTCAAGGCTGTGAGCCTGGGGCCCTGGGTTCACCTGATTG 1549
Qy 1442 AGGATCTGGCTTGT---CCCAAGAACACAGCCGGTTGAGAGCATGAAATTTGACACG 1498
Db 1550 AGGAGCTGAGCTGTTTGACAGAGCCCATGAGAGCCTGGTGTATCTCAGAGCAAGA 1609
Qy 1499 ATTGGCTCTGTGGCTCCCATACTGAGGTGACACAAGTGAACACCAAGCAACTGTGGCC 1558
Db 1610 AGCTGCTCTTTCGCGCTCCCGCTCTCAGCTGTGAGCTGCGCGTGGCCGAGTGCATAA 1669
Qy 1559 GTCTCAGAGCTGCTGGAGTGTATCTTGGCCAGGACCCCGTGTGCGCTTGAGCTTCC 1618
Db 1670 AGTATCGCTCTGTGACAGACTGTGCTCGCCGGGACCCCTATTGCGCCTGGAGGTCA 1729
Qy 1619 GCTTGTGATGCTTGTGGCCCAACAGCGGCGA 1649
Db 1730 ACACAGCCGCTGTGTGGCGTGGGTGGCCA 1760

RESULT 9

US-10-002-304-11
; Sequence 11, Application US/10002304
; Publication No. US20030036185A1
; GENERAL INFORMATION:
; APPLICANT: Shinkets, Richard
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Yang, Meljia
; APPLICANT: Boldog, Ferenc
; APPLICANT: Herrmann, John
; TITLE OF INVENTION: Polynucleotides and polypeptides encoded thereby
; FILE REFERENCE: 15966-554 Cura-54 CON-S8
; CURRENT APPLICATION NUMBER: US/10/002.304
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 09/604,286
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,584
; PRIOR FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2155
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (166)..(1935)
US-10-002-304-11

Query Match 10.7%; Score 248.6; DB 14; Length 2155;
Best Local Similarity 52.1%; Pred. No. 8e-63;
Matches 704; Conservative 0; Mismatches 629; Indels 18; Gaps 6;

QY	308	GAAGATCGACTGATGGTACTGAGACTCACAGACAGAACTGGAGGAAGAAAGCGAAG-	366
DB	419	GAGCGATCTCTGGAGGCCCCCGTGGAGAAAGAAAGACTGAGTGTATCCAGAAAGGA	478
QY	367	--AAAGAGAGCAATGTACAAATTTATCCAGATTCTCGCCATTGTCAATGCCCTCTACC	424
DB	479	ACACACGACGAGTCTCAACTTCATCCGCTTCTTGCAGCCCTACAAATGCCCTCCACC	538
QY	425	TCCTACGCTGGGACACCTTCGCTTTTATCCGAAGTGGGGGTTATTGATGTCTCAAGTT	484
DB	539	TGTACGCTGTGGCACCCTACGCCCTTCCAGCCAAAGTGCACCTAGCTCAACATGCTCACT	598
QY	485	TCCA--CGAGTTGAAGACTTGAGAGCGCGCGGGGAATGTCTTTTACGCCAGTCTC	541
DB	599	TCACTTTGGAGCATGGAGAGTTTGAAGATGGGAAGGGCAAGTGTCCCTATGACCAAGCTA	658
QY	542	AACGTCAGCAGCTGTAATGGCTGGGGCGTCCCTCTACACCGCCACTGTGAAGAACTTCC	601
DB	659	AGGGCATGCTGGCCTTCTTGTGGATGTGAGCTGTACTCGGCCACACTCAACAACTTCC	718
QY	602	TGGGACTGAGCCCATCATCTCCGAGCTGTGGTCTGAGCTGAGAGCTGGATTTCGAACAG	661
DB	719	TGGCAGCAAGAACCAATATCTCGCTTAACATGGGGC---CCACCACTCCATGAAGACAG	775
QY	662	AGACCTTCTCATCTGGCTTAATGCTCCAGCTTTGTGCGAGCTATGCTCTGAGCCAG	721
DB	776	AGTACCTGGCCTTTGGGCTCAACGAACCTCACTTTGTAGGCTCTGCCCTATGACTGAGA	835
QY	722	CTGAGTGGGGGATGAAGATGAGACGATGAAATCTTTTCTTTCACGGAGACCTGCC	781
DB	836	GTGTGGGAGCTTACGGGGGACGACGAGGTCTACTTCTTCTCAGGAGCGGGCAG	895
QY	782	GAGTGTGGACTCTTATGAGCGCATCAAGGTCCCAAGAGTGGCGGAGTGTGGCGGGG	841
DB	896	TGGAGTCCGACTGCTATCCGAGCAGAGTGGTGGCTCGTGTGGCCGCTGCTCAAGGCG	955
QY	842	ACCTTGGGGGAGGAGACCTTTCAGCAGAGATGGACGACCTTCTGAAGCTGACCTGC	901
DB	956	ATATGGGGGGCGACGGACCTTGCAGAGGAAGTGGACCACCTTCTGAAGCGCGGCTGG	1015

QY	902	TGTGCCAGGCGCCGAGCATGGCCGGGCTCCGGGGTTCTGAGGCTATGACAGGCTTC	961
DB	1016	CATGCTCTGCCCGCAACTGGCAGCTCTACTTCAACCAGCTCAGCGGATG---CACACCC	1072
QY	962	GGCCTCAGCCTGGAGCGGGAACCCCCATCTTTTATGGGATCTTTTCTCCAGTGGGAAG	1021
DB	1073	TGCAGGACACCTCTCTGGCACAACACACCTCTTTTGGGTTTTCAGACACAGTGGGTG	1132
QY	1022	GAGCTGCCATCTCTGCTGTGCTCTCCGACCCCAAGACATCGGCGCACTGCTGAATG	1081
DB	1133	ACATGTACTCTCGGCCATCTGTGAGTACCAGTTGGAAAGATCCAGCGGTTGTTGAGG	1192
QY	1082	GTCCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTCCCTGTCTATGGACAACAGG	1141
DB	1193	GCCCTATATAGGAGTACCATGAGGAAGCCCAAGAGTGGGACCGCTACACTGACCT--G	1249
QY	1142	TGCCCCAGCCAGACCTGGAGAGTGCATCGCCACAACATGAAGTCCAGCAGTTTGGAT	1201
DB	1250	TACCCAGCCCTCGGCTGGCTGATTAACAACATGGCATCGGCGCCACCGCTACACCA	1309
QY	1202	CCTCACTCTCCCTGCCAGACCGCTGCTCACTTTATCAGAGACACCCCTCTCATGGACA	1261
DB	1310	GCTCCCTGGAGCTACCCGACAACATCTCAACTTCGTCAAGAGCACCCGCTGATGGAGG	1369
QY	1262	GGCCGCTGTTCCCGGCTGACGGCGCCCTGCTGCTGCTCACTACAGATACAGCCTATCTCA	1321
DB	1370	AGCAGGTGGGCTCGGTGGAGCGCCCTGCTCGTGAAGAGGCAACCACTTCACCC	1429
QY	1322	GAGTGTGGCCACAGGCTGACACGCTCTCAGGAAAGAAATATGACGTCTTACCTGG	1381
DB	1430	ACCTGGTGGCGACCGGTTACAGACTTGATGGAGCCACCTATACAGTGTCTCATTTG	1489
QY	1382	GGACAGAGGATGGACACCTCCACCGGCTGTGCGCATTTGGAGCTCAGTGTCTTGG	1441
DB	1490	GCACAGGAGAGGCTGCTCAAGGCTGTGAGCTGGGGCCTGCGGCTTCACTGTATTG	1549
QY	1442	AGGATCTGGCCTTGT---CCAGAACACACGCGGTTTGAGAGCATGAAATTTGACCAAG	1498
DB	1550	AGGAGCTGCAGCTGTTTGACCCAGGAGCCCATGAGAAGCTTGTCTATCAGAGCAAGA	1609
QY	1499	ATTGGCTCTGGTGGGCTCCCATCTAGGTGACACAGTGAACACCAAGCACTGTGGCC	1558
DB	1610	AGCTGCTTTTGGCGGCTCCCGCTCTCAGCTGGTGCAGCTGCCCTGGCCGACTGCATAA	1669
QY	1559	GTCTCCAGAGCTGCTCGGAGTGTATCTCGGCCAGGACCCGCTGTCGCCCTGGAGCTTCC	1618
DB	1670	AGTATCGCTCTGTGCAGACTGTCTCTCGCCCGGACCCCTATTTCGCCCTGGAGCGTCA	1729
QY	1619	GGCTTGATGCTGTGTGGCCCAAGCCGCGCA	1649
DB	1730	ACACAGCGCTGTGTGGCCGTGGGTGGCCA	1760

RESULT 10

US-10-003-152-21
; Sequence 21, Application US/10003152
; Publication No. US20020151494A1
; GENERAL INFORMATION:
; APPLICANT: Shinkets, Richard
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Yang, Meljia
; APPLICANT: Boldog, Ferenc
; APPLICANT: Herrmann, John
; TITLE OF INVENTION: No. US20020151494A1el Amino Acid Sequences for Human Semaphorin
; FILE REFERENCE: 15966-554 Cura-54 CON-S12
; CURRENT APPLICATION NUMBER: US/10/003.152
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 09/604,286
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,584
; PRIOR FILING DATE: 1999-06-23

; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 2156
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (166)..(2037)
; NAME/KEY: variation
; LOCATION: (1)..(2156)
; OTHER INFORMATION: N may be any nucleotide
US-10-003-152-21

Query Match 10.7%; Score 248.6; DB 13; Length 2156;
Best Local Similarity 52.1%; Pred. No. 8e-63;
Matches 704; Conservative 0; Mismatches 629; Indels 18; Gaps 6;

QY 308 GAAGGATGACGTGGATGTAACATTTATCCAGATTCTCGCCATTGTCAATGCTCTCACC 366
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D 419 GAGCGATCTCTGGGAGGCCCGGTGGAGAAAGACTGAGTGTATCCAGAAAGGAAGA 478
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 367 --AAGAGACCAATGTACAAATTTATCCAGATTCTCGCCATTGTCAATGCTCTCACC 424
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D 479 ACAACAGACCGAGTGTCTCAACTTCATCCGCTTCTCGAGCCCTACAAATGCTCTCCACC 538
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 425 TCCTCAGCTGCGGCACCTTCGCTTTTGATCCGAAGTGGGGGTTATGTATGTGTCCAGTT 484
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D 539 TGTACGTCTGTGCGACCTTACGCTTCCAGCCCAAGTGCACCTTACATGCTCACTTCACT 598
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 485 TCCA--GCAGTTTGAAGACTTTGAGAGCGGCCGGGGAAATGTCTTTTGAAGCAGCTC 541
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D 599 TCACCTTGGAGCATGGAGATTGAAGATGGGAAGGCAAGTGTCCCTATGACCCAGCTA 658
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 542 AACGCTACGACCTGTAATGGCTGGGGGCTCTACACCCGCTTGTGAAGACTTCC 601
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D 659 AGGGCCATCTGGCCCTTCTTGTGGATGGTGTACTCTGCGCCACACTCAACAACCTTCC 718
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 602 TGGGACTGAGCCCATCTCCGAGCTGTGGTTCGAGCTGAGGACTGGATTGCAACAG 661
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D 719 TGGGACGGAACCCATATCTTCGCTTAACATGGGC---CCACCACTCCATGAAGACAG 775
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 662 AGACCTTGTACCTCTGCTTAATGCTTCAGCTTGTGCGACCTATGCTCTGAGCCGAG 721
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D 776 AGTACTGCGCTTTTGGCTCAACGAACCTCACTTTGTAGGCTCTGCTATGTACCTGAGA 835
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 722 CTGAGTGGGGGATGAAGATGAGACGATGAATCTTTTCTTCTACGGAGACCTCCC 781
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D 836 GTGTGGCGAGCTTACGGGGGACGACGAAGGTCTACTTCTTTCAGGGAGCGGGCAG 895
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 782 GAGTGTGACCTCTATGAGCGCATCAAGTCCCAAGAGTGGCCGAGTGTGTGGGGG 841
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D 896 TGGAGTCCGACTGCTATGCCGAGAGGTGGTGGCTGTGTGGCCGCTGTCTCAAGGGCG 955
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 842 ACCTTGGGGGAGGAGACCTTTCAGAGAGATGAGACGCTTCTGAGGCTGACCTGTC 901
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D 956 ATATGGGGGGCCACGACCTTGCAGAGGAAGTGGACCACTTCTCTGAAGGCGGCGCTGG 1015
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 902 TGTGCCAGGGGCCCGAGCATGCCGGGCTCCGGGTTCTCGAGGCTATGGCAGACTTC 961
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D 1016 CATGCTCTGCCCGAATGCGAGCTCTACTTCAACCAAGTGCAGGGCGATG---CACACC 1072
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 962 GGCCTCAGCTGGAGGGGAACCCCATCTTTTATGGATCTTTTCTCCAGTGGGAG 1021
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D 1073 TGCAGGACACCTCTCGGCACACACACCTTCTTTGGGGTTTTCAGCACAGTGGGTG 1132
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1022 GAGCTGCCATCTCTGCTGTGTGCTTCCGACCCCAAGACATCGGGGAGTGTGAATG 1081
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D 1133 ACATGTACCTGTGCGGCCATCTGTGAGTACCACTTGGAAAGATCCAGCGGGTGTGTTGAG 1192
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1082 GTCCCTTTAGAGACTAAACATGACTGCAACAGGGGACTGCTGTCATGGACACAGG 1141
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D 1193 GCCCCTAAGGATACCATGAGGAAGCCCAAGTGGGACCGCTACACTGACCCT---G 1249
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1142 TGCCCCAGCCAGACCTGGAGAGTGCATGCGCAACAACATGAAGTCCACAGTTTGGAT 1201
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D 1250 TACCCAGCCCTCGGCTGCGTGCATTAAACAACCTGGCATCGCGCCACGCTACACCA 1309
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1202 CCTCACTCTCCCTGCCAGACCGGCTGCTCACCTTTATCAGAGACCACTCTCATGAGGA 1261
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D 1310 GCTCCTGGAGCTACCCGACACATCTCAACTTCTGTCGAAGAGGACCGGCTGATGGAGG 1369
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1262 GGCCCGTGTCCCGCTGACGGCCGCCCTCTGCTGCTACATACAGATACAGCCTATCTCA 1321
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D 1370 AGCAGTGGGGCTCGGTGGAGCGCCCTCTCTGCTGAAGAGGACCAACTTCAACCC 1429
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1322 GAGTGTGCCCCACAGGCTGACAGCCTCTCAGGGAAGAATATGACGTGCTCTACCTGG 1381
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D 1430 ACCTGCTGCCGACCGGTTACAGGACTTGATGGAGCCACCTATACAGTGTCTTTCATTG 1489
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1382 GGACAGAGATGGACACCTCCACCGGCTGTGCGGATTTGGAGCTCAGCTCACTGTCTGG 1441
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D 1490 GCACAGGAGACGGCTGGCTGCTCAAGGCTGTGAGCTGGGGCCTGGGTTCACTGATTTG 1549
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1442 AGGATCTGSCCTTGT---CCAGAACACACACCGGTTGAGAGCATGAATTTGTACCAG 1498
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D 1550 AGGAGCTGAGCTGTTTACAGAGAGCCCATGAGAGCCTGTGTCTATCTCAGACGAAA 1609
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1499 ATTGGCTCTGTGGGCTCCCATACTAGAGGTGACACAAAGTGAACACCACTGTGGCC 1558
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D 1610 AGCTGCTCTTTCGCGCTCCCGCTCTCAGCTGTGCGAGCTGCGCGGCTGACATTA 1669
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1559 GTCTCCAGAGCTGCTCGGAGTGTATCTGGGCCAGGACCCCTGTGCGGCTGGAGCTTC 1618
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D 1670 AGTATCGCTCTGTGACAGACTGTCTCTCGCCGGGACCCCTATTGGCCTGGAGGTCA 1729
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1619 GCTTGTGCTGTGTGGCCACGCGGCGA 1649
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
D 1730 ACACAGCGCTGTGTGGCTGGTGCCCA 1760
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
US-10-002-050-21
; Sequence 21, Application US/10002050
; Publication No. US2003032095A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Yang, Meijia
; APPLICANT: Boldog, Ferenc
; APPLICANT: Herrmann, John
; TITLE OF INVENTION: No. US20030032095A1el Nucleic Acid Sequences Encoding Human Se
; FILE REFERENCE: 15966-554 Cura-54 CON-S14
; CURRENT APPLICATION NUMBER: US/10/002,050
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 09/604,286
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/140,584
; PRIOR FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 2156
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (166)..(2037)
; NAME/KEY: variation
; LOCATION: (1)..(2156)
; OTHER INFORMATION: N may be any nucleotide
US-10-002-050-21

Query Match 10.7%; Score 248.6; DB 14; Length 2156;
Best Local Similarity 52.1%; Pred. No. 8e-63;

Matches 704; Conservative 0; Mismatches 629; Indels 18; Gaps 6;			
Qy	308	GAAGGATCGACTGATGGTACCTCAGACTCAGACAGAACTGCGAGGAAGAGCAAG- 366	
Db	419	GAGGATCTCTGGAGGCCCCGTGGAGAAGAGACTGAGTGTATCCAGAAAGGGAAGA 478	
Qy	367	--AAAGAGGAGAAATGCAAAATTTATCCAGATTCGCGCAATGTCAAATGCCCTCAC 424	
Db	479	ACAAACAGAGCGAGTGTCTAACTTCATCGCTTCCTCGAGCCCTACAATGCCCTCCACC 538	
Qy	425	TCCTCAGCTGGGACCTTCGCTTTTATCCGAAGTGGGGTTATTGATGTGTCAGTT 484	
Db	539	TGTAAGCTGTGGGACCTACGCTTCAGGCCCCAAGTGCACCTACGTAACATGTCACCT 598	
Qy	485	TCCA--GCAGTTGAAGACTTGAGAGCGCGCGGGGAAATGTCCTTTTGGAGCAGCTC 541	
Db	599	TCACCTTTGGAGCATGGAGAGTTTGAAGATGGGAAGGCAAGTGTCCCTATGACCCAGTA 658	
Qy	542	AACGGTCAGACAGCTGTAATGGCTGGGGGCTCTCTACACGGCCACACTGTGAAGAACTTC 601	
Db	659	AGGCCATGCTGGGCTTCTTTGTGATGGTGAAGTGTACTCGGCCACACACTCAACAACCTCC 718	
Qy	602	TGGGAGCTGAGCCCATCATCTCCGAGCTGTGGGTCGAGCTGAGGACTGGATTGCAACAG 661	
Db	719	TGGCAGCGAAACCCATATCTCGGTAACTATGAGGCG--CCACCACTCCATGAAGACAG 775	
Qy	662	AGACCTTGTCTATCCTGGCTTAATGCTCCAGCGCTTTGTGCGAGCTATGCTCTGAGCCAG 721	
Db	776	AGTACCTGGCTTTTGGCTCAACGAACCTCACTTTGTAGGCTCTGCCTATGACTGAGA 835	
Qy	722	CTGAGTGGGGGATGAAGATGGAGACGATGAAATCTTTTCTTTCACGAGACCTTCCTCC 781	
Db	836	GTGTGGGAGCTTTCACGGGGACGACGACAAAGTCTACTTCTTTCAGGAGCGGGGAG 895	
Qy	782	GAGTGTGGACTCTATGAGGCGATCAAGTTCACAGAGTGGCGCGAGTGTGCGGGGG 841	
Db	896	TGGAGTCGAGCTATGTCGAGAGAGTGGTGTGCTGTGGCGCGGTCTGCAAGGCGC 955	
Qy	842	ACCTTGGGGGAGGAAGACCTTCAGCAGAGATGAGCAGCGCTTTCTGAAGCTGACCTGC 901	
Db	956	ATATGGGGGGGACGACCTGCAGAGGAAGTGGACCACTTCCTGAAGGCGGCTGG 1015	
Qy	902	TGTCGCCAGGCGCGAGTGGCGGGCTCCGGGGTTCGCGAGGCTATGCGCAGAGTTC 961	
Db	1016	CATGCTGTGCCCCGAACCTGGCAGCTCTACTTCAACCACTGCGAGCGATG---CACACCC 1072	
Qy	962	GGCTCAGCTGGAGCGGAACCCCATCTTTTATGGATCTTTTCTCCAGTGGGAAG 1021	
Db	1073	TGCAGGACACCTCTGGCACACACACACCTTCTTTGGGTTTTCAGCAGAGTGGGTG 1132	
Qy	1022	GAGTGCCTATCTGCTGTGTGCTTCCGACCCCAAGACATCCGGGCGAGTGTGAATG 1081	
Db	1133	ACATGTACCTGTGCGCATCTGTGAGTACCACTGTTGGAAGAGATCCAGCGGTGTTGAGG 1192	
Qy	1082	GTCCCTTTAGAGACTAAACATGACTGCAACAGGGGACTGCCTGTATGAGCAACAGG 1141	
Db	1193	GCCCTATAAGAGATACCATGAGGAAGCCCAAGAGTGGGACCGGTACACTGACCTC---G 1249	
Qy	1142	TGCCCCAGCCCGAGACCTGGAGAGTGCATCGCCCAACATGAAGCTCCAGCAGTTGGAT 1201	
Db	1250	TACCCAGCCCTGCGGCTGGCTGTGCTATTAACAACTGGCATCGGCGGCCCGGTACACCA 1309	
Qy	1202	CCTCAGCTCTCCCTGCCAGACCGCTGTCTACCTTTATCAGAGACACCTCTCATGGACA 1261	
Db	1310	GCTCCCTGGAGCTACCCGACAACATCTCACTTTCGTCAAGAGCACCCGCTGATGGAG 1369	
Qy	1262	GGCCCGTGTCCCGGCTGACGGCGCCCTGTGCTGTCATCAGATACAGACCTTATCTCA 1321	
Db	1370	AGCAGGTGGGGCTCTGGTGGAGCGCGCCCTGCTCGTGAAGAGGGACCAACTTCACCC 1429	
Qy	1322	GAGTGTGGCCCCAGAGGTGACCACTCTCAGGGAAGAAATATACGCTGCTCTACCTGG 1381	
Db	1430	ACCTGGTGGCGGACCGGGTTACAGGACTTGATGGAGCCACCTATACAGTGTGTTCAATG 1489	

Qy	1382	GGACAGAGATGGACAACTCCACCGGCTGTGCGCATTTGGAGCTCAGCTCAGTGTCTGG 1441	
Db	1490	GCACAGGAGAGGCTGCTCAAGGCTGTGAGCCTGGGCGCTTGGTTTACCTGATTG 1549	
Qy	1442	AGGATCTGGCCTTTGTT---CCAGAACACAGCCGGTTGAGAGCATGAATTTGACCAAG 1498	
Db	1550	AGGAGCTGCAGCTGTTTGACCAAGAGCCCATGAGAAGCCTGGTCTATCTCAGAGCAAA 1609	
Qy	1499	ATTGCGTCTGGTGGCTCCCATACTGAGGTGACACAAAGTGAACACCAAGCAACTGTGCC 1558	
Db	1610	AGCTGCTCTTTGGCGGCTCCCGCTCAGCTGTGTCAGCTGCCGTGGCCGACTGCATTA 1669	
Qy	1559	GTCTCCAGAGCTGCTCGGAGTGTATCTGCGCCAGGACCCCGTGTGCGCTGGAGCTTCC 1618	
Db	1670	AGTATCGCTCTGTGACAGCTGTCTCTCGCCGGGACCCCTATTGGCGCTGGAGCGTCA 1729	
Qy	1619	GGCTGTGATGCTGTGTGGCCCAACCGCGCA 1649	
Db	1730	ACACAGCGGCTGTGTGGCGCTGGGTGGCCA 1760	

RESULT 12

US-10-002-304-21
; Sequence 21, Application US/10002304
; Publication No. US20030036185A1
; GENERAL INFORMATION:
; APPLICANT: Shlmkels, Richard
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Yang, Meljia
; APPLICANT: Boldog, Ferenc
; APPLICANT: Herrmann, John
; TITLE OF INVENTION: Polynucleotides and polypeptides encoded thereby
; FILE REFERENCE: 15966-554 Cura-54 CON-S8
; CURRENT APPLICATION NUMBER: US/10/002.304
; PRIOR FILING DATE: 2001-11-02
; PRIORITY APPLICATION NUMBER: 09/604,286
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,584
; PRIOR FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 21
; LENGTH: 2156
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (166)..(2037)
; NAME/KEY: variation
; LOCATION: (1)..(2156)
; OTHER INFORMATION: N may be any nucleotide
US-10-002-304-21

Query Match 10.7%; Score 248.6; DB 14; Length 2156;			
Best Local Similarity 52.1%; Pred. No. 8e-63;			
Matches 704; Conservative 0; Mismatches 629; Indels 18; Gaps 6;			
Qy	308	GAAGGATCGACTGATGGTACCTCAGACTCAGACAGAACTGCGAGGAAGAGCAAG- 366	
Db	419	GAGGATCTCTGGAGGCGCCCGTGGAGAAGAGACTGAGTGTATCCAGAAAGGAAGA 478	
Qy	367	--AAAGAGGAGAAATGTCACAAATTTATCCAGATTCGCGCAATTTGTCATATGCCTCTCACC 424	
Db	479	ACAAACAGAGCGAGTGTTCACCTTCATCCGCTTCCTGCGACCCCTACAAATGCCCTCCACC 538	
Qy	425	TCCTCAGCTGGGCACTTCGCTTTTGTATCCGAAGTGGGGTTTATTGATGTGTCCAGTT 484	
Db	539	TGTAGCTGTGGCACCTTACGCTTCAGAGCCCAAGTGCACCTACGTCAACATGCTCACT 598	
Qy	485	TCCA---GCAGTTGAAGAGACTTGAGAGCGCGCGGGGAAATGCTCTTTTGAGCGCAGCTC 541	

Db 599 TCACCTTTGAGCATGGAGAGTTTGAAGATGGGAAGGCAAGTGTCCCTATGACCAGCTA 658
QY 542 AACGGTCAGACAGCTGTAATGGCTGGGGCGTCTCTACACGCCCACTGTGAAGAAGTCTCC 601
Db 659 AGGGCCATGCTGGCCCTTCTTGTGATGTTGAGCTGTACTCGGCCACACTCAACAAGTCTCC 718
QY 602 TGGGAGCTGAGCCCATCATCTCCCGAGCTGTGGTCGAGCTGAGGACTGGATTTCGAACAG 661
Db 719 TGGGCACGGAACCCATATCTCGCTGAACATGGGCG---CCACCACTCCATGAAGACAG 775
QY 662 AGACCTTGTATCTCGGCTTAATGCTCCAGCCTTTGTGCGAGCTATGTCCTGAGCCAG 721
Db 776 AGTACTCGCCCTTTTGGCTACAGAACCTCACTTTGTAGGCTCGCTATGTATACCTGAGA 835
QY 722 CTGAGTGGGGGATGAAGATGGAGAGATGAATCTTTTCTTCAAGGAGACCTCC 781
Db 836 GTGTGGGAGCTTACGGGGGACGACAGAGTCTACTTCTTTCAGGGAGCGGGCAG 895
QY 782 GAGTGTGAGCTCTATGAGCGCATCAAGGTCCCAAGAGTGGCCGAGTGTGTGGGGG 841
Db 896 TGGAGTCCGACTGCTATCCGAGCAGGTGGTGGCTCGTGTGGCCGCTGTCTCAAGGGCG 955
QY 842 ACCTTGGGGGAGGAGACCTTTCAGCAGAGATGACAGACGTTTCTGAAGCTGACCTGC 901
Db 956 ATATGGGGGGCACCGGACCCCTGCAGAGGAAGTGCACACGTTCTGAAGCGCGGCTGG 1015
QY 902 TGTGCCCGGCGGCGAGCATGGCGGGCTCGGGGTTCTGCAGGCTATGGCAGAGCTTC 961
Db 1016 CATGCTCTGCCCGAAGTGGCAGCTCTACTTCAACCAAGCTGCAGCGGATG---CACACCC 1072
QY 962 GGCCTCAGCTGGAGCGGGAACCCCATCTTTTATGGATCTTTTCTCCAGTGGGAG 1021
Db 1073 TGCAGGACACCTCTCTGGCACAACACCACTTCTTTGGGGTTTTTCAAGCACAGTGGGTTG 1132
QY 1022 GAGCTGCCATCTCTCTGTGTGTGCTTCCGACCCCAAGACATCGGGCAGTGTGAATG 1081
Db 1133 AGATGACTCTGCGGCCATCTGTGAGTACCAAGTGGGAAGATCAGCGGGTGTGTGAGG 1192
QY 1082 GTCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCTGTCTATGACACAGG 1141
Db 1193 GCCCTATAGAGGTACCATGAGGAGGCCAGAGTGGGACCGCTACACTGACCT---G 1249
QY 1142 TGCCTCAGCCAGACTGGAGAGTGCATCGCCAAACATGAAGCTCCAGCAGTTTGGAT 1201
Db 1250 TACCCAGCCCTCGGCTGCTGCTGATTAACAACTGGCATCGGGCGCCACGGCTACACCA 1309
QY 1202 CCTCACTCTCTCGCCAGCCCGTGTCACTTTTACAGAGACCACTCTCATGGACA 1261
Db 1310 GCTCCTGGAGTACCCGACACATCTCTCAACTTCGTCAAGAGCACCCGCTGATGGAGG 1369
QY 1262 GCGCGTGTTCGGGCTGACGGCGCCCGCTGCTGCTCACTACAGATACAGCCTATCTCA 1321
Db 1370 AGCAGTGGGCTCGGTGGAGCCCGCCCTGCTGCTGAAGNAGGGCACTTCAACC 1429
QY 1322 GAGTCGTGGCCACAGGTGACACGCTCTCAGGAAAGAAATATGACGTGCTTACCTGG 1381
Db 1430 ACCTGGTGGCCACCGGTTACAGGACTTGTGAGGACCACTATACAGTGTGTTCAATTG 1489
QY 1382 GGACAGAGATGGACACCTCCACCGGCTGTGCGCATTTGGAGCTCAGCTCAGTGTCTGG 1441
Db 1490 GCACAGGAGAGCGGTGGCTGCTCAAGGCTGTGAGCTGGGCGCTGACCTGATG 1549
QY 1442 AGGATCTGGCCTTGT---CCAGAACCAACAGCCGTTTGAGAGCATGAATTTACACAG 1498
Db 1550 AGGAGCTGAGCTGTTTGACCAAGAGCCCATGAGAAGCTTGTGTATCTCAGAGCAAA 1609
QY 1499 ATTGGCTCGTGGGTCCCATACTGAGGTGACACAAGTGAACACCAAGCACTGTGGCC 1558
Db 1610 AGCTGCTCTTTGGCGGCTCCCGCTCTCAGCTGGTGCAGCTGCGCGGCGCACTCATTA 1669
QY 1559 GTCTCCAGAGCTGCTCGAGTGTATCTTGGCCAGGACCCGCTGTGGCGCTGGAGCTTCC 1618
Db 1670 AGTATCGCTCTGTGCAGACTGTGTCTTCGCGCGGAGCCCTTATTGGCCCTGGAGCGTCA 1729

QY 1619 GCGTTGATCTGTGTGGCCCAACCGCGGCA 1649
Db 1730 ACACCAGCGCTGTGTGGCGGTGGTGCCA 1760

RESULT 13

US-10-003-152-13

; Sequence 13, Application US/10003152

; Publication No. US20020151494A1

; GENERAL INFORMATION:

; APPLICANT: Shimkets, Richard

; APPLICANT: Fernandes, Elma

; APPLICANT: Vernet, Corine

; APPLICANT: Yang, Meijia

; APPLICANT: Boldog, Ferenc

; APPLICANT: Herrmann, John

; TITLE OF INVENTION: No. US20020151494A1el Amino Acid Sequences for Human Semaphori

; FILE REFERENCE: 15966-554 Cura-54 CON-S12

; CURRENT APPLICATION NUMBER: US/10/003.152

; CURRENT FILING DATE: 2001-11-02

; PRIOR APPLICATION NUMBER: 09/604, 286

; PRIOR FILING DATE: 2000-06-22

; PRIOR APPLICATION NUMBER: 60/140, 584

; PRIOR FILING DATE: 1999-06-23

; NUMBER OF SEQ ID NOS: 49

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 13

; LENGTH: 2284

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (166)..(1953)

; NAME/KEY: Variation

; LOCATION: (1)..(2284)

; OTHER INFORMATION: N may be any nucleotide

US-10-003-152-13

Query Match 10.7%; Score 248.6; DB 13; Length 2284;

Best Local Similarity 52.1%; Pred. No. 8.2e-63;

Matches 704; Conservative 0; Mismatches 62; Indels 18; Gaps 6;

QY 308 GAAGGATCAGCTGGATGCTACCTTGAGCTCAGACAGACAACTGCAGGAAGAAAGCAAG- 366
Db 419 GAGCGATCTCTGGGAGGCCCGCTGGAGAAAGAACTGAGTGTATCCAGAAAGGAAGA 478
QY 367 --AAGAGGAGATGTCAAAATTTTATCCAGATTCTCGCCATTTGCAATGCTCTCACC 424
Db 479 ACAACGACGCGAGTGTCAACTTCATCCGCTTCTCGACGCCCTACAAATGCTCCCAACC 538
QY 425 TCCTCAGCTGGGCACTTCGCTTTTATCGAAGTGGGGTTATTGATGTGTCCAGTT 484
Db 539 TGTAGCTGTGGCACTTACCGCTTCCAGCCCAAGTGCACCTACATGCTCACT 598
QY 485 TCCA---CGAGTTGAAAGACTTGAGAGCGCGGGGAAATGTCCTTTTGAAGCAGCTC 541
Db 599 TCACCTTGGAGCATGGAGATTTGAAGATGGAGGGCAAGTGTCCCTATGACCCAGCTA 658
QY 542 AACGTCAGCAGCTGTAATGCTGGGGCGTCTCTACACGCCCACTGTGAGAACTTCC 601
Db 659 AGGCGCATGCTGGCCTTCTTGTGATGCTGAGCTACTCGGCCACACTCAACAACCTTCC 718
QY 602 TGGGACTGAGCCCATCATCTCCCGAGCTGTGGGTCGAGCTGAGCACTGGATTTCGAACAG 661
Db 719 TGGGCACGGAACCCATTTATCTGCTAACATGGGCG---CCACCACTCCATGAAGACAG 775
QY 662 AGACCTTGTATCTCGGCTTAATGCTCCAGCCTTTGTGCGAGCTATGTCCTGAGCCAG 721
Db 776 AGTACCTGGCCTTTTGGCTCAACGAACCTCACTTTGTAGGCTCTGCCTATGTACCTGAGA 835
QY 722 CTGAGTGGGGGATGAAGATGGAGAGCATGAATCTTTTCTTCTACGGAGACCTCCC 781

Db 836 GTGTGGGAGCTTCACGGGGACGACGACAGGCTTACTTCTTTCAGGGAGCGGCAG 895
Qy 782 GAGTGTGGACTCTATGAGCGCATCAAGTTCACAGAGTGGCCCGAGTGTGTCGGGG 841
Db 896 TGGAGTCCGACTGCTATGCGGAGCAGGTGTGGCTGTGGCCCGTGTCTGCAAGGCG 955
Qy 842 ACCTTGGGGGAGAAACCCCTTCAGCAGAGATGGACGAGCTTTCCTGAAGCTGACCTGC 901
Db 956 ATATGGGGGGCGACGACCCCTGCAGAGAGTGGACCAAGTTCCTGAAGCGCGGCTGG 1015
Qy 902 TGTGCCAGGCGCGAGCATGGCGGGCTCCGGGTTCTGACAGCTATGCGAGAGCTTC 961
Db 1016 CATGCTGTGCCGGAATCGGAGCTTACTTCAACAGCTGCAGGCGATG---CACACC 1072
Qy 962 GGCTCAGCTGGAGCGGGAAACCCCATCTTTTATGGATCTTTTCTCCAGTGGGAAG 1021
Db 1073 TGCAGGACACTCTCTGCGCAACACACCACTTCTTTGGGTTTTCAGACAGTGGGGTG 1132
Qy 1022 GAGTGGCATCTCTGTGTGTGTGCTTCGACCCCAAGACATCCGGGCGAGTGTGAATG 1081
Db 1133 ACATGTACCTGTGGCCATCTGTGAGTACCAGTTTGAAGAGATCCAGCGGTGTGTTGAG 1192
Qy 1082 GTCCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCCTGTCTATGCACACAGG 1141
Db 1193 GCCCTATAGGAGTACCATGAGGAGCCCAAGAGTGGGACCGTACACTGACCT---G 1249
Qy 1142 TGCCCCAGCCGACCTGGAGAGTGCATCGCCAAACATGAAGCTCCAGCAGTTGGAT 1201
Db 1250 TACCAGCCCTCGGCCCTGCTGTCATTAACAACTGGCATCGCGCCACGGCTACACCA 1309
Qy 1202 CTTCACTCTCCCTGCCAGAGCGGTGCTACCTTTATCAGAGACCCCTCTCATGAGCA 1261
Db 1310 GCTCCTCTGGAGTACCCGCAACATCTCAACTTCGTAAGAAGCAACCCGCTGATGAGG 1369
Qy 1262 GGCCCGTGTCCGGCTGACGGCGCCCTGCTGTGTCACACTACAGATACAGCTATCTCA 1321
Db 1370 AGCAGTGGGGCTCTGGTGGAGCGCCCTGCTGCTGTAAGAAGGGACCACTTCACCC 1429
Qy 1322 GAGTGTGGCCCCAGGGTGACCAAGCTCTCAGGGAAGAAATATGAGTGTCTTACCTGG 1381
Db 1430 ACCTGTGGCGGACCGGTTACAGGACTTGTATGAGCCACCTATACAGTGTGTCATTG 1489
Qy 1382 GGACAGGATGACACACTCCAGCGGCTGTGGCATTTGAGCTGACGTACGTGTGTTGG 1441
Db 1490 GCACAGAGAGCGCTGGCTCTCAAGGCTGTGAGCCCTGGGGCCCTTGGGTTTCACTGATTG 1549
Qy 1442 AGGATCTGGCTTGT---CCGAGAACACAGCCGGTTGAGAGCATGAATTTGTACCAG 1498
Db 1550 AGGAGTGCAGCTGTTTGACAGGAGGCCATGAGAAGCTGTGTCTATCTCAGAGCAAGA 1609
Qy 1499 ATTTGGCTCTGGTGGCTCCCTACTAGAGTGCACAAAGTGAACACCAAGCACTGTGGCC 1558
Db 1610 AGCTGTCTTTGCGGCTCCGCTCTCAGCTGTGCGAGCTGCCGTGGCGGAGTGCATAA 1669
Qy 1559 GTCTCAGAGCTGCTGGAGTGTATCTGTGCCAGGACCCGCTGTGGGCGCTGAGGCTTC 1618
Db 1670 AGTATCGCTCTGTGAGAGTGTGCTCTCGCCGGGACCCCTATTGGCGCTGGAGCGTCA 1729
Qy 1619 GGCTTGATGCTGTGTGGCCACCGCGGCA 1649
Db 1730 ACACAGCCGCTGTGTGGCGCTGGTGGCCA 1760

RESULT 14

US-10-002-050-13
; Sequence 13, Application US/10002050
; Publication No. US20030032095A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Yang, Meijia
; APPLICANT: Boldog, Ferenc

; APPLICANT: Herimann, John
; TITLE OF INVENTION: No. US20030032095A1el Nucleic Acid Sequences Encoding Human Se
; FILE REFERENCE: 15966-554 Cura-54 CON-S14
; CURRENT APPLICATION NUMBER: US/10/002,050
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 09/604,286
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,584
; PRIOR FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 13
; LENGTH: 2284
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (166)..(1953)
; NAME/KEY: variation
; LOCATION: (1)..(2284)
; OTHER INFORMATION: N may be any nucleotide
US-10-002-050-13

Query Match 10.7%; Score 248.6; DB 14; Length 2284;
Best Local Similarity 52.1%; Pred. No. 8.2e-63;
Matches 704; Conservative 0; Mismatches 629; Indels 18; Gaps 6;

Qy 308 GAAGGATCGACTGGATGCTACCTGAGACTCACACAGAACTGCAGGAAGAAAGCGCAAG- 366
Db 419 GAGCGATCTCTGGGAGCGCCCGTGGAGAAAGAACTGATGTATCAGAAAGGGAAGA 478
Qy 367 --AAAGAGGACGAATGTCACAAATTTATCCAGATTCCTCGCATTTGTAATGCCCTCTCAC 424
Db 479 ACACACAGCGAGTGTCTCAACTTCATCCGCTTCCTGCAGCCCTACATAGCTTCCACCC 538
Qy 425 TCCTCAGCTGGGACACCTTCGCTTTTGTATCCGAAGTGGGGGTTATTTGATGTGTCCAGTT 484
Db 539 TGTACGCTCTGTGGCACCTACGCCCTTCCAGCCCAAGTGCACCTACGTCAACATGCTCACCT 598
Qy 485 TCCA---GCAGTTTGAAGACTTGAGAGCGCGCGGGGAATGTCCTTTTTCAGCGCACTC 541
Db 599 TCAGTTTGGAGATGGAGAGTTTGAAGATGGGAAGGGAAGTGTCCCTATGACCCAGCTA 658
Qy 542 AACGTCAGCAGCTGTAAATGGCTGGGGCGCTCTACACCGCCACTGTGAAGAACTTCC 601
Db 659 AGGCCATCTGGCCCTTCTTGTGATGCTGAGCTGTACTCGGCCACACTCAACAACCTCC 718
Qy 602 TGGGAGTACGCCCATCATCTCCGAGCTGTGGGTGAGCTGAGGACTGGATTGCAACAG 661
Db 719 TGGGACGGAACCCATTTATCTGCGTAACATGGGGC---CCACCACTTCCATGAAGACAG 775
Qy 662 AGACCTTGTCTATCTGCTTAACTCTCAGCCCTTGTGCGAGCTATGCTCTGAGCCAG 721
Db 776 AGTACCTGGCTTTTGGCTCAACGAACCTCACTTTGTAGGCTTGTGCTATGATACCTGAGA 835
Qy 722 CTGAGTGGGGGATGAAGATGGAGACGATGAAATCTTTTTCCTCAGGAGACCTTCC 781
Db 836 GTGTGGCAGCTTCACGGGGAGCAGCAGAGGTCTACTTCTTCTCAGGAGGCGGCGAG 895
Qy 782 GAGTGTGGACTCTCTATGAGCGCATCAAGTTCACAGAGTGGGCCGAGTGTGTGCGGGGG 841
Db 896 TGGAGTCCGACTGTGTCGCGAGAGGTGTGGCTGTGGCCGCTGTCTGCAAGGGCG 955
Qy 842 ACCTTGGGGCAGGAAGACCTTTCAGCAGAGATGGAGAGCTTCTGAAGGCTGACCTGC 901
Db 956 ATATGGGGGGCGCAGGACCTTCAGAGGAAGTGGACAGCTTCTCCTGAAGCGCGCTGG 1015
Qy 902 TGTGCCAGGGCCCGAGCATGGCGGGCTCTCCGGGTTCTGCAAGCTATGCGAGAGCTTC 961
Db 1016 CATGCTCTGCCCGAAGTGTGCACTTCAACAGCTGCAAGGCGATG---CACACCC 1072
Qy 962 GGCTCAGCCTGAGGGGGAAACCCCATCTTTTATGGGATCTTTTCTCCAGTGGGAAG 1021

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OW nucleic - nucleic search, using sw model

Run on: September 25, 2003, 17:51:04 ; Search time 571.81 Seconds
(without alignments)
11004.333 Million cell updates/sec

Title: US-09-284-180A-2
Perfect score: 2331
Sequence: 1 atgttgccaggccgagc.....gtgatgagacctcatctaa 2331

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-n-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Query Match	Length DB ID	Description	
1	2331	100.0	2331 19	AAV07280	
2	2331	100.0	4008 19	AAV07279	
3	2138.8	91.8	4024 21	AAZ87977	
4	1827.6	78.4	2971 22	AAF93776	
5	1626	69.8	2893 20	AAK86126	
6	1626	69.8	2894 20	AAK86127	
7	1393.2	59.8	2315 19	AAV07281	
8	1391.2	59.7	1761 19	AAV07282	

9	654	28.1	799	20	AAK86128	EST sequence for D
10	567.6	24.4	928	25	ABT32087	NOVX DNA sequence
11	527.4	22.6	777	25	ABT32088	NOVX DNA sequence
12	505.2	21.7	671	24	ABL89791	Human polynucleoti
13	372.4	16.0	669	22	AAF93972	Primer specific fo
14	251.8	10.8	3293	22	AAD08048	Human extracellular
15	248.6	10.7	2155	22	AAC84887	Human SEC6 nucleic
16	248.6	10.7	2156	22	AAC84892	Human SEC11 nucleic
17	248.6	10.7	2284	22	AAC84888	Human SEC7 nucleic
18	248.6	10.7	3556	22	AAD08283	Human secreted pro
19	248.6	10.7	3776	24	ABN83983	Human gene sequenc
20	247.8	10.6	2558	24	AAD28948	Human MOL5b cDNA.
21	247.8	10.6	3112	24	AAD28949	Human MOL5c cDNA.
22	247	10.6	3781	21	AAA37092	Human PRO1480 (UNQ
23	247	10.6	3781	22	AAS46151	Human DNA encoding
24	247	10.6	3781	22	AAF54381	Primer #75 used in
25	247	10.6	3781	25	ACAS7909	Human PRO1480 cDNA
26	247	10.6	3781	25	ABX98379	Human cDNA encodin
27	247	10.6	3781	25	ABX98881	Novel human secret
28	247	10.6	3781	25	ACA05926	Human secreted/tra
29	247	10.6	3781	25	ABX97970	Human PRO polynuci
30	247	10.6	3781	25	ABX78754	Human PRO polynuci
31	247	10.6	3781	25	ABX75767	Human cDNA encodin
32	247	10.6	3781	25	ABX76972	Human PRO polynuci
33	247	10.6	3781	25	ABX16812	Human cDNA encodin
34	246	10.6	3868	24	AAD28947	Human MOL5a cDNA.
35	245.4	10.5	3503	22	AAF29461	Murine M-Sema-F cD
36	245.4	10.5	3766	24	ABV77918	Hypoxia-induced pr
37	245.4	10.5	3766	24	ABN83984	Human gene sequenc
38	245.4	10.5	3766	25	ACC51059	Human bladder canc
39	245.4	10.5	3766	25	ABX76369	Lung cancer-associ
40	238.4	10.2	8095	24	ABX92031	Lung specific nucle
41	238.4	10.2	8144	25	ABT13397	Breast specific re
42	238.2	10.2	2703	21	AAA47437	Sequence encoding
43	238	10.2	333	19	AAV07286	Human semaphorin W
44	228	9.8	2390	24	ABK11109	DNA encoding human
45	224.4	9.6	2615	24	ABQ99272	Human coding sequ

ALIGNMENTS

RESULT 1
AAV07280
ID AAV07280 standard; cDNA to mRNA; 2331 BP.
XX AAV07280;
AC AAV07280;
XX
DT 08-SEP-1998 (first entry)
XX
DE Rat semaphorin W encoding cDNA.
XX
KW Rat; semaphorin W; nerve extension inhibitor; antiallergic; anticancer;
KW immunosuppressant; gene therapy; diagnosis; research reagent; ds.
XX
OS Rattus norvegicus.
XX
FH Key Location/Qualifiers
FT CDS 1..2331
FT FT /*tag= a
FT FT /product= "semaphorin W"
XX
PN W09815628-A1.
XX
PD 16-APR-1998.
XX
XX 03-OCT-1997; 97WO-JP03549.
XX
PR 09-OCT-1996; 96JP-0287636.
XX (SUMU) SUMITOMO PHARM CO LTD.
XX Kikuchi K, Kimura T;
PI

Db 796 GCTGAGTGGGGGATGAAGATGAGACGATGAATCTTTTTTTTCTTACGAGACCTCC 855
QY 781 CGAGTGTGGACTCCTATGAGCGCATCAAGGTCCCAAGATGGCCCGAGTGTGTGCGGG 840
Db 856 CGAGTGTGGACTCCTATGAGCGCATCAAGGTCCCAAGATGGCCCGAGTGTGTGCGGG 915
QY 841 GACCTTGGGGGAGGAAGACCTTTCAGCAGAGATGGACGACCTTTCTGAAGGCTGACCTG 900
Db 916 GACCTTGGGGGAGGAAGACCTTTCAGCAGAGATGGACGACCTTTCTGAAGGCTGACCTG 975
QY 901 CTGTGCCAGGCGCCGAGCATGGCCGGCTCCGGGTTCAGGCTATGCGACAGCTT 960
Db 976 CTGTGCCAGGCGCCGAGCATGGCCGGCTCCGGGTTCAGGCTATGCGACAGCTT 1035
QY 961 CGGCTCTAGGCTGGAGCGGAACCCCATCTTTTATGGGATCTTTTCTCCAGTGGGAA 1020
Db 1036 CGGCTCTAGGCTGGAGCGGAACCCCATCTTTTATGGGATCTTTTCTCCAGTGGGAA 1095
QY 1021 GGAGCTGCCATCTGTGCTGTGTGCTTCCGACCCCAAGACATCCGGGCACTGCTGAAT 1080
Db 1096 GGAGCTGCCATCTGTGCTGTGTGCTTCCGACCCCAAGACATCCGGGCACTGCTGAAT 1155
QY 1081 GGTCCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCCCTGTCATGACACAGAG 1140
Db 1156 GGTCCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCCCTGTCATGACACAGAG 1215
QY 1141 GTGCCCGCCAGACCTGGAGAGTGCATCGCAACAAACATGAAGCTCCAGAGTTTGA 1200
Db 1216 GTGCCCGCCAGACCTGGAGAGTGCATCGCAACAAACATGAAGCTCCAGAGTTTGA 1275
QY 1201 TCCTCACTCTCCCTGCCAGACCGGTGTCTACCTTTTATCAGAGACACCCCTCTCATGGAC 1260
Db 1276 TCCTCACTCTCCCTGCCAGACCGGTGTCTACCTTTTATCAGAGACACCCCTCTCATGGAC 1335
QY 1261 AGGCCCGTGTCCCGGCTGACGGCGCCCGCTGCTGCTACTACAGATACAGCTTATCTC 1320
Db 1336 AGGCCCGTGTCCCGGCTGACGGCGCCCGCTGCTGCTACTACAGATACAGCTTATCTC 1395
QY 1321 AGAGTCGTGGCCACAGAGGTGACCAAGCTCTCAGGGAAGAAATATGAGTGTCTACCTG 1380
Db 1396 AGAGTCGTGGCCACAGAGGTGACCAAGCTCTCAGGGAAGAAATATGAGTGTCTACCTG 1455
QY 1381 GGGACAGAGATGGACACTTCACCGGGGTGTGCGCATTTGGAGCTCAGTCACTGCTCTTG 1440
Db 1456 GGGACAGAGATGGACACTTCACCGGGGTGTGCGCATTTGGAGCTCAGTCACTGCTCTTG 1515
QY 1441 GAGGATCTGGCTTCTTCCAGAACACAGCGGTTGAGAGCATGAAATTTACACAGAT 1500
Db 1516 GAGGATCTGGCTTCTTCCAGAACACAGCGGTTGAGAGCATGAAATTTACACAGAT 1575
QY 1501 TGGCTCCTGGTGGGCTCCCATACTGAGGTGACACAAGTGAACACCAAGCTGTGGCGGT 1560
Db 1576 TGGCTCCTGGTGGGCTCCCATACTGAGGTGACACAAGTGAACACCAAGCTGTGGCGGT 1635
QY 1561 CTCACAGAGTGTCTGGAGTGTATCTTGGCCAGGACCCCGTGTGCGCTGGAGCTTCCGG 1620
Db 1636 CTCACAGAGTGTCTGGAGTGTATCTTGGCCAGGACCCCGTGTGCGCTGGAGCTTCCGG 1695
QY 1621 CTGTGATGTGTGTGGCCACCGCCGCGGAGCACCAGCGGATGGTTCAAGATATAGATCA 1680
Db 1696 CTGTGATGTGTGTGGCCACCGCCGCGGAGCACCAGCGGATGGTTCAAGATATAGATCA 1755
QY 1681 GCGGATGTCTCTCTTTTGTGTCAAAGAAACCTGGAGAACATCCCGTAGTGTGTTGAAGTT 1740
Db 1756 GCGGATGTCTCTCTTTTGTGTCAAAGAAACCTGGAGAACATCCCGTAGTGTGTTGAAGTT 1815
QY 1741 CCGGTGGCTACTGTGGGCGCAGTGTCTGTCATGTTCCCGAGTTCTGCTGGGCATCC 1800
Db 1816 CCGGTGGCTACTGTGGGCGCAGTGTCTGTCATGTTCCCGAGTTCTGCTGGGCATCC 1875
QY 1801 TGTGTGTGGCACCACCGCCAGTGGATGACTGGCTCACTCCCGGAGGATGGACTAGAG 1860
Db 1876 TGTGTGTGGCACCACCGCCAGTGGATGACTGGCTCACTCCCGGAGGATGGACTAGAG 1935

QY 1861 GTGGTGGTGAACCCAGGGGCCATGGGGGCTTATGCTTGGAGTGTCAAGAGGTGGAGCC 1920
Db 1936 GTGGTGGTGAACCCAGGGGCCATGGGGGCTTATGCTTGGAGTGTCAAGAGGTGGAGCC 1995
QY 1921 GCGCCGCTGGTGGCTGCTTATAGCTTGTGTGGGCGACCGACGGGGACCTCAAAACGG 1980
Db 1996 GCGCCGCTGGTGGCTGCTTATAGCTTGTGTGGGCGACCGACGGGGACCTCAAAACGG 2055
QY 1981 GCGCACACCGTGTGGGGCTGATTTGGTTGGCTTTCCTCTGGGTGTTCTTCACACATCC 2040
Db 2056 GCGCACACCGTGTGGGGCTGATTTGGTTGGCTTTCCTCTGGGTGTTCTTCACACATCC 2115
QY 2041 CTCACCTCTCCTCTGATTGGTGCCTCAGCAGCGTTCGCGCAGACAGAGGAGCTTCTAGCT 2100
Db 2116 CTCACCTCTCCTCTGATTGGTGCCTCAGCAGCGTTCGCGCAGACAGAGGAGCTTCTAGCT 2175
QY 2101 AGAGACAAGTGGGCTTAGATCTGGGGCTCCACCTTCTGGGACCAACAGCTATAGTCAG 2160
Db 2176 AGAGACAAGTGGGCTTAGATCTGGGGCTCCACCTTCTGGGACCAACAGCTATAGTCAG 2235
QY 2161 GACCCCTCCCTCTCCTTCCGCTGAAGATGAACGCTGCCCTCGCCCTGGGTAAAGCGGCG 2220
Db 2236 GACCCCTCCCTCTCCTTCCGCTGAAGATGAACGCTGCCCTCGCCCTGGGTAAAGCGGCG 2295
QY 2221 AGTGGTTTTGTGGCTTCCCTCCACCTTCTCCTGCTGGATTCTTGCCCAAGCCAGCCAC 2280
Db 2296 AGTGGTTTTGTGGCTTCCCTCCACCTTCTCCTGCTGGATTCTTGCCCAAGCCAGCCAC 2355
QY 2281 ATCCGGCTCACTGGGGCGCTCTAGCCAGCTGTGATGAGACCTCCATCTAA 2331
Db 2356 ATCCGGCTCACTGGGGCGCTCTAGCCAGCTGTGATGAGACCTCCATCTAA 2406

RESULT 3

AAZ87977
ID AAZ87977 standard; DNA; 4024 BP.

XX AAZ87977;

DT 05-JUN-2000 (first entry)

XX Mouse semaphorin W polypeptide encoding DNA.

XX Semaphorin W gene; hereditary disease; Parkinson's disease; mouse; ds.

XX Mus musculus.

XX Key Location/Qualifiers

FT CDS 79..2412

FT /*tag= a

XX WO200006725-A1.

XX 10-FEB-2000.

XX 30-JUL-1999; 99WO-JP04120.

XX 31-JUL-1998; 98JP-0217467.

XX 24-FEB-1999; 99JP-0046833.

XX (SUMU) SUNITOMO PHARM CO LTD.

XX Kimura T, Encinas JA;

XX WPI; 2000-195293/17.

XX P-PSDB; AAY77803.

XX Diagnosis of hereditary diseases such as Parkinson's disease by
PT analyzing the occurrence of mutation in the semaphorin W gene, or
PT determining the genotype of the semaphorin W gene in an individual

XX Disclosure; Page 22-28; 40pp; Japanese.


```
QY 1978 CGGGCCACACCGTTGTGGGGCTGGATTGGTTGGCTTTCTCTGGGTGTTCTTTCAGCA 2037
Db |||||
QY 2059 CGGGCCACACCGTTGTGGGGCTGGCTGGTTGGCTTTTCTCTGGGTGTTCTTTCAGCA 2118
Db |||||
QY 2038 TCCTCAGCTCTCTCTCTGATTGCTGCGCTCAGCAGCGTGGGCGACAGAGGAGCTTCTA 2097
Db |||||
QY 2119 TCCTCAGCTCTCTCTCTGATTGCTGCGCTCAGCAGCGTGGGCGACAGAGGAGCTTCTA 2178
Db |||||
QY 2098 GCTAGAGACAAGGTGGGCTTAGATCTGGGGGCTTCCACCTTCTGGGACCAACAAGCTATAGT 2157
Db |||||
QY 2179 GCTAGAGACAAGGTGGGCTTAGATCTGGGGGCTTCCACCTTCTGGGACCAACAAGCTATAGC 2238
Db |||||
QY 2158 CAGGACCCCTCTCTCTGCTGAAGATGAACGCTGCCCTGCCCTGGTGAAGCGG 2217
Db |||||
QY 2239 CAAGACCCCTCTCTCTCCTTCACTGAAGATGAACGCTGCCCTGGCTGGTGAAGCGG 2298
Db |||||
QY 2218 GGCAGTGGTTTGGTGGCTTCCCTCCACCTTCTCTGCTGATTCTTGGCCCAAGCCAGCC 2277
Db |||||
QY 2299 GGCAGTGGTTTGGTGGCTTCCCTCCACCTTCTCTGCTGATTCTTGGCCGAGCCAGCC 2358
Db |||||
QY 2278 CACATCCGGCTCAGTGGGCGGCTCTAGCCACGCTGTGATGAGACCTTCCATCTAA 2331
Db |||||
QY 2359 CACATCCGGCTCAGTGGGCGGCTCTAGCCACGCTGTGATGAGACGCTTCTATCTAA 2412

RESULT 4
AAF93776
ID AAF93776 standard; cDNA; 2971 bp.
XX
AC AAF93776;
XX
AT 23-MAY-2001 (first entry)
XX
DE Human cDNA encoding a membrane or secretory protein clone PSEC0074.
XX
KW Human; secretory protein; membrane protein; vaccine; gene therapy;
KW rheumatoid arthritis; diabetes; ss.
XX
OS Homo sapiens.
XX
PN EP1067182-A2.
XX
PD 10-JAN-2001.
XX
PF 07-JUL-2000; 2000BP-0114090.
XX
PR 08-JUL-1999; 99JP-0194179.
PR 11-JAN-2000; 2000JP-0118775.
PR 02-MAY-2000; 2000JP-0183766.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;
XX
DR WPI; 2001-093989/11.
DR P-PSDB; AAB88349.
XX
PT Nucleic acids encoding secretory proteins/membrane proteins, useful in
PT gene therapy or as candidate target molecules in drug development -
XX
PS Claim 1; SEQ ID 65; 609pp + CD ROM; English.
XX
CC This invention relates to nucleic acid sequences AAF93744 - AAF93916
CC AAB88317 - AAB88419. Included in the invention are primers
CC AAF93917 - AAF94295 and AAF62232 - AAF62235 which are used to isolate the
CC cDNA sequences of the invention. The invention also includes methods for
CC the production of antibodies directed against the proteins, and cDNA
CC sequences, which can be used in vaccines. The polynucleotide sequences
CC can be used in gene therapy. The polynucleotide sequences and the
CC proteins they encode may be used in the prevention, treatment and
CC diagnosis of diseases associated with inappropriate secretory
CC protein/membrane protein expression. The nucleic acids and complementary
```

```
CC sequences may also be used as DNA probes in diagnostic assays
CC (e.g. polymerase chain reactions (PCR)) to detect and quantitate the
CC presence of similar nucleic acid sequences in samples. They may also be
CC used to study the expression and function of secretory proteins/membrane
CC polypeptides and their role in metabolism. The polypeptides may be used
CC as antigens in the production of antibodies against them and in assays to
CC identify modulators (agonists and antagonists) of expression and
CC activity. The antibodies and antagonists may also be used as therapeutic
CC agents to down regulate expression and activity. The antibodies may also
CC be used as diagnostic agents for detecting the presence of the
CC polypeptides in samples (e.g. by enzyme linked immunosorbant assay
CC (ELISA)). Examples of diseases which may be treated include rheumatoid
CC arthritis and diabetes.
XX
SQ Sequence 2971 bp; 582 A; 853 C; 858 G; 678 T; 0 other;
```

```
Query Match 78.4%; Score 1827.6; DB 22; Length 2971;
Best Local Similarity 87.3%; Pred. No. 0;
Matches 2033; Conservative 0; Mismatches 279; Indels 18; Gaps 2;

QY 1 ATGCTTGGCCAGGCGCGCCGCCCGCCCGCGCGCGCTCCGGTCTTTCCTTC 60
Db |||||
QY 103 ATGCGGCGCTCTGCTGCGGCGCGCGCGCGCGCTCCGGGCGAGCTACAGCTCGCCCTTC 162
Db |||||
QY 61 CCGCGCGCGCTCTGCTGCTGCTGCTGGCGATACCTAAGCGCCCGGCTGGCGCGC 120
Db |||||
QY 121 GTCCCGCGCTAGTCCCGAGAACCTCGCTGCCATCTCCGAGGCTACCTCTATCTCACC 180
Db |||||
QY 208 GTCCCGCGCTCGGTGCCCGAGAACCTCGCTTCCAATCTCTGAGGCTGACTCTCTCACC 267
QY 181 CGGTTTGGACGCTCTCATACGTACAAATTACTCTGCTCTCTCTTGTGATCTCTGCCCTCCAC 240
Db |||||
QY 268 CGGTTGCGAGTCCCTCACACATACAAATTACTCTGCTCTCTTGTGATCTCTGCCCTCCAC 327
QY 241 ACACCTTACGTCGGTGCACGGGATAGCATCTTCGGCTTTAACCTCCTCCCTTCTTGGGAA 300
Db |||||
QY 328 ACACCTTATGTTGGGCGCGCGGACACCATCTTCGCTTTATCGCTGCGCTTCTCAGGGGAG 387
QY 301 AGACCCCGAAGATCGACTGGATGTACCTGAGACTCACAGACTCACAGACAGAACTGAGGAGAA 360
Db |||||
QY 388 AGACCCCGAGGATTGACTGGATGGTTCCTGAGGCTCACAGACAGAACTGAGGAGAA 447
QY 361 GGCAAGAAAGAGAGCAATGTACAAATTTATTCAGATTTCTGCCATTTCTCAATGCTCT 420
Db |||||
QY 448 GGCAAGAAAGAGAGCAATGTACAAATTTGTTCAGATTTCTGCCATTTCCCAATGCTCT 507
QY 421 CACCTCTCACGTCGGGCGACCTTCGCTTTTGTATCCGAAGTGGGGGTTTATGATGTGCC 480
Db |||||
QY 508 CACCTCTCACGTTGTGGCACCCTTCGCTTTTGTATCCGAAGTGGGGGTTTATGATGTGCC 567
QY 481 AGTTTCCAGCAGGTTGAAGACTTCAGAGCGCGCGGGGAAATGTCTTTTTCAGCCAGCT 540
Db |||||
QY 568 AGTTTCCAGCAGGTTGAAGACTTCAGAGTGGCGGGGAAATGTCTTTTTCAGCCAGCT 627
QY 541 CAACGCTCAGCAGCTGTAAATGGCTGGGCGCTCTCTACACCGCCACTGTGAAGAACTTC 600
Db |||||
QY 628 CAGCGCTCAGCAGCTGTAAATGGCTGGGCGGCTCTCTATGCTGCCACTGTGAAGAACTAC 687
QY 601 CTGGGGAGTACGCCCATCATCTCCGAGCTGTGGGTGAGCTGAGAGCTGGATTGGAACA 660
Db |||||
QY 688 CTGGGGAGGAGCAATATCACCAGACGAGTGGGTGCTGGCGAGGACTGGATTGCGACA 747
QY 661 GAGACTTGTCTATCTGCTTAACTCTCAGGCTTTGTGCGAGCTATGGTCTCTGAGCCCA 720
Db |||||
QY 748 GATACCTTGGCTTCTCTGCTGACGCCCGGCTTTGCGGACCGCTGGCCCTTGAGCCCA 807
QY 721 GCTGAGTGGGGGATGAAGATGGAGACGATGAATCTTTTTCCTTTCAGGAGACCTCC 780
Db |||||
QY 808 GCCGAATGGGGGATGAAGATGGAGACGACGAAATCTACTTCTTCTTTCAGGAGACTCC 867
QY 781 CGAGTGTGGACTCTCTATGAGGCGCATCAAGGTCCCAAGAGTGGCCCGGAGTGTGCGGGG 840
```


CC and detection of a mutation in its polynucleotide may also be used to
CC diagnose a disease or condition or susceptibility to a disease related
CC to altered expression or activity of SBSEMN1. These diseases and
CC conditions include neurodegeneration, spinal injury, neuropathies,
CC neuromuscular disorders, muscular dystrophy, psychiatric disorders,
CC inflammatory disorders, developmental malformations, disorders of the
CC immune system, cancer and viral infections. The protein may also be
CC used to treat these diseases and conditions by administration as a
CC vaccine.

SQ Sequence 2893 BP; 579 A; 825 C; 842 G; 647 T; 0 other;

Query Match 69.8%; Score 1626; DB 20; Length 2893;

Best Local Similarity 83.2%; Pred. No. 0;

Matches 1938; Conservative 0; Mismatches 275; Indels 117; Gaps 3;

QY 1 ATGCTTGCCAGGGCCGAGCGGCCCGCCCGGGCCCCCGGGCTTTCCCTTC 60

nb
110 ATGCGGGCTGTGCCTCCGCCCGCCCGCAGCCCTACACCGTCCCCCTTC
168

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

22 170 CC
GCATCTGCTGCTTGGCGGTGCTGAGCGGCCCGGTAACCGCCCGC Z14

121 GTCCCGGCTCAGTGCCGAGAACCTCGCTGCCCATCTCCGAGGCTGACTCCTATCTCACC 180

Db 215 GTCCCCCGCTCGGTGCCAGAACCTCGCTTCCAAATCTCTGAGGCTGACTCCTGTCTCACC 274

QY 181 CGGTTTGCAGCGTCTCATACGTACAAATTACTCTGCTCTCCTTGTGGATCCTGCCCTCCAC 240

Db 275 CGGTTGCAGTCCCTCACACATACAAATTACTCTGTTCTCCTTGTGGATCCTGCCCTCCCAC 334

[illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

.....

QY JVI AGACCCGAAAGGACIGGATGGTACCTGAGACICACAGACAGAAACIGCAGGAAGAAA 360

DB 393 AGACCCCGCAGGATTTGACTGGATGGTTCCCTGAGGCTCACAGACAGAACTGTAGGAAGAAA 454

361 GGCAAGAAAGAGGACGAATGTCACAAATTTATCCAGATTCTGCCAATTGTCAATGCCCTCT 420

Db 455 GGCAAGAAAGAG----- 466

QY 421 CACCTCCTCACGTGCGGCACCTTCGCTTTTGATCCGAAGTCCGGGTTATTGATGTGTCC 480

db 467 475
-----CATGTCCTCC

401

[illegible]

QY 341 CAACGGTCAAGCAGCTGTAAATGGCTGGGGGGCTCTCTACACGCCACCTGTGAAGAATTC 600

Db 536 CAGCGGTCAGCAGCTGTAATGGCTGGGGGGTCTCTATGCTGCCACTGTGAAACTAC 595

QY 601 CTGGGACTGAGCCCATCATCTCCGAGCTGTGGGTCGAGCTGAGACTGGATTCTGAACA 660

Db 596 CTGGGGACGGAGCCAAATTATCACCAGAGCAGTGGGTCTGTCGGAGGACTGGATTCTGGACA 655

661 GAGACCTTCTCAATCCCTCGCCTTAATTGCTCCACCCCTTTCTTCGCCAATCCCTCATCCCCA

[illegible]

QY /ZI GCIGAGTGGGGGAGTGAAGATGGAGACGATGAAATCTTTTCTTCACGGAGACCTCC /80

Db 716 GCCGAATGGGGGATGAAGATGGAGACGACGAAATCTACTTCTTACGGAGACTTC 775

781 CGAGTGTGGACTCCTATGAGCGCATCAAGGTCCCAAGAGTGCGCCCGAGTGTGTGCGGGG 840 QY

Db 776 CGAGCATTTGACTCATACGAGCGCATTAAGTCCACGGTGGCCCGTGTGTGCGGG 835

QY 841 GACCTTGGGGGCAGGAAGACCCCTTCAGCAGAGATGGACGACGTTTCTGAAGGCTGACCTG 900

[illegible]


```
Db 836 GACCTCGGGGCGGGAAGACCTCCAGCAGAGATGGACGACGCTTTTGAAGACTGACCTG 895
QY 901 CTGTGCCAGGCGCCGAGCATGCGGGGCTCCGGGGTTCTCAGGCTATGCAGAGCTT 960
Db 896 CTCTGTCCAGGCGCTGAGCATGGCGGGCTCCAGTGTCTCCAGGATGTTGCTGTGCTT 955
QY 961 CGGCTCAGCCTGGAGCGGGGAACCCCATCTTTTATGGGATCTTTTCTCCAGTGGGAA 1020
Db 956 CGACCTGAGCTTGGGCGAGGGAATCCCATCTTTTATGGCATCTTTTCTTCCAGTGGGAG 1015
QY 1021 GGAGTGGCATCTCTGCTGTGTGTGCTTCCGACCCCAAGACATCGGGCAGTGTGAAT 1080
Db 1016 GGGGTACTATCTCTGCTGTGTGCTTCCGACCAAGACATTCGGACAGTGTGAAT 1075
QY 1081 GGTCCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCCCTGTATGACACACGAG 1140
Db 1076 GGTCCCTTCAGAGAACTAAACATGACTGCAACAGGAGACTGCCCTGTGTCGTGACAAATGAT 1135
QY 1141 GTGCCCGCCAGACACTGGAGAGTGCATGCGCAACACATGAAGTCTCAGCAGTTTGA 1200
Db 1136 GTGCCCGCCAGACACTGGAGAGTGCATGACCAACAACATGAAGTCTCGGCACCTTTGGC 1195
QY 1201 TCTCTACTCTCCCTGCGAGACCGCTGCTCACTTTATCAGAGACACCCCTCTCATGGAC 1260
Db 1196 TCATCTCTCTCCCTGACCGCGTACTCACTTCACTCCGGGACCAACCACCTCATGGAC 1255
QY 1261 AGGCCCGTGTCCCGGCTGACGGCGGCCCTGCTGGTCACTACAGATACAGCCTATCTC 1320
Db 1256 AGGCCAGTGTTCACAGCTGATGCCCCACCCCTGCTGCTCACTACAGATACAGCCTATCTC 1315
QY 1321 AGAGTGTGGGCCACAGGCTGACACGCTCTCAGGGAAGAATATGACGTGCTTACCTG 1380
Db 1316 AGAGTGTGGGCCACAGGCTGACACGCTCTCAGGGAAGAATATGATGCTGCTTACCTG 1375
QY 1381 GGGCAGAGAGTGGACACTCCACCGGCTGCGCATTTGGAGCTCAGTCACTGCTCTG 1440
Db 1436 GAAGATCTGGCCTTATCCAGAGCCACAGCCAGTTGAGAACAATGTTACACACG 1495
QY 1501 TGGCTCTGGTGGGCTCCCATCTAGGTGACAAAGTGACACCAAGCAACTGTGGCGT 1560
Db 1496 TGGCTCTGGTGGGCTCCCATCTAGGTGACAAAGTGAAATACAAACCAACTGTGGCGT 1555
QY 1561 CTCACAGAGTGTCTGGAGTGTATCTGCGCCAGGACCCGCTGTCGCGCTGGAGCTTCCGG 1620
Db 1556 CTCACAGAGTGTCTAGAGTGCATCTGGCCCGAGGACCCAGTCTGTGCTGGAGCTTCCGG 1615
QY 1621 CTGTATGCTGTGTGGCCACCGCGGAGCACCCGCGGATGGTTCAAGATATAGAGTCA 1680
Db 1616 CTGGATGAGTGTGTGGCCCATGSCCGGGAGCACCCAGGGTTGGTCCAAGACATAGATCA 1675
QY 1681 GCGGATGTCTCTTTGTGTCCAAAGAACCTCGGAGAACATCCCGTAGTGTGTTGAAGTT 1740
Db 1676 GCAGATGTCTCTCTTTGTGTCTTAAGAGCCGTGGAGAACGTCCTAGTAGTGTGTTGAAGTT 1735
QY 1741 CCGGTGGCTACTGTGGCCACGCTGCTGCGCATCTCCCGCAGTCTGCTGGCGATCC 1800
Db 1736 CCGGTGGCTACTGTGGCATGTGTGTGCGATGTTCTCCAAGCTCAGCATGGGCATCC 1795
QY 1801 TGTGTGTGGCACCGCCAGCTGGAGTGACTGCGCTCACTCCCGGAGGGATGGAGTAG 1860
Db 1796 TGTGTGTGGCACCGCCAGCTGGAGTGACTGCACTCACCCCGCGGGATGGAGTAG 1855
QY 1861 GTGGTGTGACCCAGGGCCATGGGGGCTTATGCTTGGAGTGTCAAGAGGTGGAGCC 1920
Db 1856 GTGGTGTGACCCAGGGCCATGGGGGCTTATGCTTGTGATGTCAAGAGGTGGGCA 1915
QY 1921 GCCCGCTGTGGCTGTCTATAGCTTGTGTGGGCGACGAGCGGGACCCCTCAACCGG 1980
Db 1916 GCCCATGTGGTAGCAGCTTACAGCTTGGTATGGGGCGACCCAGCAGATGCTCCGAGCCGG 1975

RESULT 7
AAV07281
ID AAV07281 standard; cDNA to mRNA; 2315 BP.
XX
AC AAV07281;
XX
DT 08-SEP-1998 (first entry)
XX
DE Human semaphorin W encoding cDNA with 3'UTR.
XX
KW Human; semaphorin W; nerve extension inhibitor; antiallergic; anticancer; immunosuppressant; gene therapy; diagnosis; research reagent; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1764
FT FT /*tag= a
FT FT /product= "semaphorin W"
FT FT 1765..2315
FT FT /*tag= b
XX
PN WO9815628-A1.
XX
PD 16-APR-1998.
XX
PF 03-OCT-1997; 97WO-JP03549.
XX
PR 09-OCT-1996; 96JP-0287636.
XX
PA (SUMU ) SUMITOMO PHARM CO LTD.
XX
PI Kikuchi K, Kimura T;
XX
DR WPI; 1998-261015/23.
DR P-PSDB; AAW51314.
XX
PT Nerve extension inhibitor protein semaphorin W - is useful as therapeutic drug and diagnostic and research reagent
XX
PS Claim 2; Page 65-66; 90pp; Japanese.
XX
CC The present sequence encodes human semaphorin W. Semaphorin W and its derivatives are nerve extension inhibitors which are useful as antiallergic, immunosuppressant and anticancer agents. The DNA encoding semaphorin W can also be used in gene therapy, e.g. using a viral vector. The proteins, peptides, DNA and antibodies which
```


CC recognise the protein or peptides, can be used as diagnostic or
CC research reagents. Semaphorin W can be used as a screen for
CC semaphorin W antagonists with possible therapeutic use.

XX
SQ Sequence 2315 BP; 463 A; 640 C; 670 G; 542 T; 0 other;

Query Match 59.8%; Score 1393.2; DB 19; Length 2315;
Best Local Similarity 87.2%; Pred. No. 0;
Matches 1540; Conservative 0; Mismatches 223; Indels 3; Gaps 1;

Qy	565	GGGGCGCTCTACACCGCCACCTGTAAGAACTTCCTGGGACTTGAGCCCATCATCTCC	624
Db	1	GGGGGTGCTCTATGCTGCCACCTGTGAATACCTTACCTGGGAGGAGCAATATACCC	60
Qy	625	CGAGCTGTGGTGCAGCTGAGGACTTGAATCGAACAGAGACTTGTCTATCTGGCTTAAT	684
Db	61	AGAGCAGTGGCTGTGGCGGAGGACTGGATTTCGGACAGATACCTTGCTCTGCTGAAC	120
Qy	685	GCTCAGCCTTTGTCGAGCTATGCTCTGAGCCAGCTGAGTGGGGGATGAAGATGGA	744
Db	121	CCCCCAGCCTTTGTCGACGCCGTGGCCTTGAGCCAGCCGAATGGGGGATGAAGATGGA	180
Qy	745	GACGATGAATCTTTTCTTTCACGGAGACCTCCCGAGTGTGGACTCTCTATGAGCGC	804
Db	181	GACGACGAATCTACTCTCTTACGGAGACTTCCCGAGGACTTGAATCATAGAGCGC	240
Qy	805	ATCAAGTCCCAAGAGTGGCCCGAGTGTGCGGGGACCTTGGGGGAGGAAGACCCCTT	864
Db	241	ATTAAGTCCACAGGTGGCCCGTGTGTGCGGGGACCTCGGGGCGGAGACCTC	300
Qy	865	CAGCAGATGGAGCAGCTTTCTGAAGCTGACCTGTGTGCGCAGGCCCGAGATGGC	924
Db	301	CAGCAGATGGAGCAGCTTTTGAAGCTGACCTGTCTGTCCAGGGCCCTGAGCATGGC	360
Qy	925	CGGCTCCGGGTTCTCAGGCTATGCGAGCTTGGCCCTCAGCTCGGAGCGGAACC	984
Db	361	CGGCTCCAGTGTCTCAGGATGTTCTGTGTCTGACCTTGAGCTTGGGGGAGGACT	420
Qy	985	CCCATCTTTTATGGGATCTTTTCTCCAGTGGGAAGGAGCTGCCATCTCTGCTGTGT	1044
Db	421	CCCATCTTTTATGGCATCTTTTCTCCAGTGGGAGGGGCTACTATCTCTGCTGT	480
Qy	1045	GCCTTCGACCCCAAGACATCGGGCAGTGTGAATGTGCTCCCTTTAGAGCTTAAACAT	1104
Db	481	GCCTTCGACCCACAGACATTCGGACAGTGTGAATGTGCTCCCTTCAGAGAACTAAACAT	540
Qy	1105	GACTGCAACAGGGACTGCTGTATGACACAGAGTGGCCAGCCAGACCTGGAGAG	1164
Db	541	GACTGCAACAGAGACTGCTGTATGACACATGATGTGCCCCAGCCAGACCTGGAGAG	600
Qy	1165	TGCATCGCCAAACATGAAGCTCCAGCAGTTTGGATCCTCACTCTCCCTGCCAGACCGC	1224
Db	601	TGCATCAACCAACATGAAGCTCCGGACCTTTGGCTCATCTCTCCCTGCCAGACCGC	660
Qy	1225	GTGCTCACCCTTATCAGAGACCCCTCTCATGACAGCGCCGTTCCTCCGGCTGACGGC	1284
Db	661	GTACTCACCCTTATCCTCGGAGACCCCACTCATGACAGCGCCAGTGTTCAGCTGATGGC	720
Qy	1285	CGCCCTCTGCTGCTACATACAGATACGCTATCTCAGAGTCTGGCCCAAGGGTGACC	1344
Db	721	CACCCCTCTGCTGCTACTACAGATACAGCTTATCTCAGAGTCTGGCCCAAGGGTGACC	780
Qy	1345	AGCCTCTCAGGAAAGATATGAGCTGTCTACCTGGGACAGAGATGGACACCTCCAC	1404
Db	781	AGCCTCTCAGGAAAGATATGATGTCTCTACCTGGGACAGAGATGGACACCTCCAC	840
Qy	1405	CGGCTGTGCGCATTTGGAGCTCAGCTAGTGTCTTTGGAGATCTGGCCTTGTTCACGAA	1464
Db	841	CGAGCAGTGGCGATCGGAGCTCAGCTCAGCTGCTTGAAGATCTGGCCTTATTCACAG	900
Qy	1465	CCACAGCCGTTGAGAGCATGAATTTGACACGATTTGGCTCTGGTGGGCTCCCATACT	1524
Db	901	CCACAGCCAGTTGAGAACATGAATTTGACACAGCTGGCTCTGTTGGTCCCGTACT	960

Qy	1525	GAGGTGACACAAGTGAACACACAGCAACTGTGTGCCCTCTCCAGAGCTGCTCGAGTGTATC	1584
Db	961	GAGGTGACACAAGTGAATACAACCAACTGTGTGCCCTCTCCAGAGCTGCTCAGAGTGATC	1020
Qy	1585	CTGGCCACAGGACCCCGTGTGCCCTGGAGCTTCCGGCTGTGATGTGTGGCCACAGCC	1644
Db	1021	CTGGCCACAGGACCCCACTGTGTGCCCTGGAGCTTCCGGCTGGATGATGTGTGGCCATGC	1080
Qy	1645	GGCAGACACCGGGATGTTCAAGATATAGAGTCAGCGGATGTCTCTTTTGTGTCCA	1704
Db	1081	GGGAGACACCGAGGGTGTGTCAGACATAGAGTCAGCAGATGTCTCTTTTGTGTCT	1140
Qy	1705	AAAGAACCTGGAGAACATCCCGTAGTGTGAAAGTTCGGGTGGCTACTGTGTGGCCACGTG	1764
Db	1141	AAAGAGCCTGGAGAACGTCACAGTAGTGTGAAAGTTCGGGTGGCTACTGTGTGGCCATGTG	1200
Qy	1765	GTCTGCCATGTTCGCCCAGTTCCTGCTGGGATCCTGTGTGTGGCACACAGCCAGTGA	1824
Db	1201	GTCTTGCCATGTTCCTCAAGCTCAGCATGTGTAAGTTCGGGTGGCTACTGTGTGGCCATGTG	1260
Qy	1825	GTGACTGCGCTCACCTCCCGGAGGATGGACTAGAGTGTGGTGCACCCAGGCGCATATG	1884
Db	1261	GTGACTGACATCACCCCGGGGATGGACTGTGAGTGTGTGTGACCCAGGCGCATATG	1320
Qy	1885	GGGCTTATGCTTGCAGTGTGAGAGGTGGAGCGCCGCGTGTGTGTGTATAGC	1944
Db	1321	GGGCTTATGCTGTGAATGTGAGAGGTGGGCGAGCCCATGTGTGTAGCAGCTTACAGC	1380
Qy	1945	TTGTTGTGGGAGGACCGGAGGACCTCAAAACCGGGCCACACCGTTGTGGGGCTGGA	2004
Db	1381	TTGTTATGGGAGGACCGAGAGATGCTCCGAGCCGGCGCCACAC---AGTGGGGCGGGA	1437
Qy	2005	TTGTTGTGCTTCTCTGSGTGTCTTCAGCATCCCTCACTCTCTCTCTGATTGTGCGC	2064
Db	1438	CTGGCTGGCTTCTCTGGGATCTCGCAGCATCCCTCACTCTCTCTCTGATTGTGCGG	1497
Qy	2065	CGTCAGCAGCGTGGCGACAGAGGAGCTTCTAGCTAGAGACAAGTGGGCTTAGATCTG	2124
Db	1498	CGTCAGCAGCAGCGCGACAGAGGAACTTCTGGCTAGAGACAAGTGGGCTTGACCTG	1557
Qy	2125	GGGGCTCACCTTTTGGGACCAACAGCTATAGTCAAGACCTCCCTCTCTCTGGCCTGAA	2184
Db	1558	GGGGCTCACCTTTTGGGACCAACAGCTATAGTCAAGACCTCCCTCTCTCTCTCTCTGAA	1617
Qy	2185	GATGAACGGCTGCCCTTGGCTTGGTAAGCGGGGACGTGTTTGGTGGCTTCCCTCCA	2244
Db	1618	GATGAACGGTGGCGCTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT	1677
Qy	2245	CCCTTCTCTGCTGATTTTGGCCAAAGCCAGCCACATCCGGCTCACTGGGGCGCTCTA	2304
Db	1678	CCCTTCTCTGCTGATTTTGGCCAAAGCCAGCCACATCCGGCTCACTGGGGCGCTCTA	1737
Qy	2305	GCCAGCTGTGATGAGACCTCCATCTA	2330
Db	1738	GCCACATGTGATGAACATCCATCTA	1763

RESULT 8

AAV07282

ID AAV07282 standard; cDNA to mRNA; 1761 BP.

XX

AC AAV07282;

XX

DT 08-SEP-1998 (first entry)

XX

DE Human semaphorin W encoding cDNA.

XX Human; semaphorin W; nerve extension inhibitor; antiallergic; anticancer;
KW immunosuppressant; gene therapy; diagnosis; research reagent; ds.

XX Homo sapiens.

OS

XX

FH	Key	Location/Qualifiers	
FT	CDS	1..1764	
FT		/*tag= a	
FT		/product= "semaphorin w"	
XX	XX	WO9815628-A1.	
XX	XX	16-APR-1998.	
XX	XX	03-OCT-1997; 97WO-JP03549.	
PF	XX	09-OCT-1996; 96JP-0287636.	
XX	XX	(SUMU) SUMITOMO PHARM CO LTD.	
XX	XX	Kikuchi K, Kimura T;	
PI	XX	WPI; 1998-261015/23.	
DR	XX	P-PSDB; AAW51314.	
XX	XX	Nerve extension inhibitor protein semaphorin W - is useful as therapeutic drug and diagnostic and research reagent	
XX	XX	Claim 2; Page 67-68; 90pp; Japanese.	
XX	XX	The present sequence encodes human semaphorin W. Semaphorin W and its derivatives are nerve extension inhibitors which are useful as anti-allergic, immunosuppressant and anticancer agents. The DNA encoding semaphorin W can also be used in gene therapy, e.g. using a viral vector. The proteins, peptides, DNA and antibodies which recognise the protein or peptides, can be used as diagnostic or research reagents. Semaphorin W can be used as a screen for semaphorin W antagonists with possible therapeutic use.	
XX	XX	Sequence 1761 BP; 349 A; 511 C; 521 G; 380 T; 0 other;	
SQ	Query Match	59.7%; Score 1391.2; DB 19; Length 1761;	
	Best Local Similarity	87.2%; Pred. No. 0;	
	Matches 1538; Conservative	0; Mismatches 223; Indels 3; Gaps 1;	
QY	565	GGGGCGTCTCTACACCCCACTGTGAAGAACTTCTCGGGAGTGGCCCATCATCTCC 624	
DB	1	GGGGGTGCTCTATGCTGCCACTGTGAAGAACTACCTGGGACGGCAATTATCACC 60	
QY	625	CGAGTGTGGGTGAGCTGAGACTGGATTTCGAACAGAGACCTTGTATCTCTGGCTTAAT 684	
DB	61	AGAGCAGTGGGTGCTGCCAGGACTGGATTTCGACAGATACCTTGCCTTCTCTGGTGAAC 120	
QY	685	GCTCAGCCTTTGTGCGACGCTATGTCCTGAGCCAGCTGAGTGGGGGATGAAGATGGA 744	
DB	121	GCCCCAGCCTTTGTGCGACGCTGGCTTGAGCCAGCCGAATGGGGGATGAAGATGGA 180	
QY	745	GACGATGAATCTTTTCTTTCAGGAGACCTCCCGAGTGTGGACTCTCTATGAGCGC 804	
DB	181	GACGAGAAATCTACTTCTTTACGGAGACTTCCCGAGCATTTGACTCATACGAGCGC 240	
QY	805	ATCAAGTCCCAAGAGTGGCCCGAGTGTGGGGGACCTTGGGGCAGGAAGACCCCTT 864	
DB	241	ATTAAAGTCCCAAGGTTGGCCCGTGTGTGGGGGGACCTCGGGGGCGGAAGACCCCTC 300	
QY	865	CAGCAGAGATGGACGACGTTTCTGAAGGCTGACCTGCTGTGCCAGGGCCCGAGCATGGC 924	
DB	301	CAGCAGAGATGGACGACGTTTGTGAAGCTGACCTGTCTGTCCAGGGCCCTGAGCATGGC 360	
QY	925	CGGGCCTCGGGTCTTCAGGCTATGGCAGAGCTTCGGCCTCAGCCTGGAGCGGGAACC 984	
DB	361	CGGGCCTCCAGTGTCTCGAGGATGTGCTGTGCTTCGACCTGAGCTTGGGCGAGGACT 420	
QY	985	CCCATCTTTATGGGATCTTTCTCCAGTGGGGAAGAGCTGCCATCTCTGCTGTGTGT 1044	
DB	421	CCCATCTTTATGGGATCTTTCTTCCAGTGGGAGGGGCTACTATCTCTGCTGTGTGT 480	
QY	1045	GCCTTCCGACCCCAAGACATCCGGGCGAGTGTGTAATGGTCCCTTTAGAGAGCTAAACAT 1104	

DB	481	GCCTTCCGACCAAGACATTCGGACAGTGTCTGAATGGTCCCTTCAGAGAACTAAACAT 540	
QY	1105	GACTGCAACAGGGGACTGCTGTATGGCAACAGAGTGGCCCGCCAGCCAGACCTGGAGAG 1164	
DB	541	GACTGCAACAGAGAGTGTCTGTGGCAATGATGTGCCCGCCAGCCAGACCTGGAGAG 600	
QY	1165	TGCATCGCCCAACATGAAGTCCAGAGTGTGGATCTCTACTCTCCCTGCCAGACCGC 1224	
DB	601	TGCATCACCACCAACATGAAGTCCCGGCACTTTGGCTCATCTCTCTCCCTGCCAGCGC 660	
QY	1225	GTGCTCACCTTTATCAGAGACCACTCTCATGCGAGGCCCCGTGTTCGGGGTGAAGGC 1284	
DB	661	GTACTCACCTTTCATCGGGGACCACTCATGAGAGGCCAGTGTTCAGCTGATGGC 720	
QY	1285	CGCCCTCTGCTGCTCACTACAGATACAGCTATCTCAGAGTGTGGCCCGCCAGAGGGTACC 1344	
DB	721	CACCCCTCTGCTGCTCACTACAGATACAGCTATCTCAGAGTGTGGCCCGCCAGAGGGTACC 780	
QY	1345	AGCCTCTCAGGAAAGAATATGAGCTGCTTACCTGGGACAGAGAGTGGACACCTCCAC 1404	
DB	781	AGCCTCTCAGGAAAGAGTATGATGTGCTTACCTGGGACAGAGAGTGGACACCTCCAC 840	
QY	1405	CGGGCTGTGCGCATTTGGAGCTCAGCTCAGTGTCTTGGAGATCTGGCTTTTCCAGAA 1464	
DB	841	CGAGCAGTGGGATCGGAGCTCAGCTCAGCGTCTTTGAAGATCTGGCTTATTTCCAGAG 900	
QY	1465	CCACAGCCGTTGAGAGCATGAAATGTACACGATTGCTCTGCTGGCTGCCATCT 1524	
DB	901	CCACAGCCAGTTGAGAACATGAAATGTACACAGCTGGCTCTGCTGGCTCCCGTACT 960	
QY	1525	GAGGTGACACAAAGTGAACACCACTGTGGCCGCTCTCAGAGCTGCTCGGAGTGTATC 1584	
DB	961	GAGGTGACACAAAGTGAATACCACTGTGGCCGCTCTCAGAGCTGCTCAGAGTGCATC 1020	
QY	1585	CTGGCCAGGACCCCTGTGGCGTGGAGCTTCCGGCTTGTAGTGTGTGGCCCGACGC 1644	
DB	1021	CTGGCCAGGACCCAGTGTGCTGTGGAGCTTCCGGCTGGATGTGTGGCCCGTGC 1080	
QY	1645	GCGAGCAGCCGGGATGGTTCAAGATATAGAGTACGCGGATGCTCTCTTTGTCTCCA 1704	
DB	1081	GGGGAGCAGCGAGGTTGGTCCAAGACATAGAGTACAGAGATGTCTCTCTTTGTGCTCT 1140	
QY	1705	AAAGAACCTGGAGAACATCCCGTAGTGTGTTGAAGTTCGGTGGCTACTGTGGGCCACGTG 1764	
DB	1141	AAAGAGCCTGGAGAACCTCCAGTAGTGTGTTGAAGTTCGGTGGCTACAGCTGCGCATGTG 1200	
QY	1765	GTCTTGCCATGTTCCTCCAGTGTGCTGGGATCTGTGTGGCCAGCCAGCCAGTGA 1824	
DB	1201	GTCTTGCCATGTTCCTCCAGTGTGCTGGGATCTGTGTGGCCAGCCAGCCAGTGA 1260	
QY	1825	GTGACTGCGCTCACCTCCCGGAGGATGGACTAGAGTGGTGGTACCCCGAGGGGCGCATG 1884	
DB	1261	GTGACTGCGCTCACCTCCCGGAGGATGGACTAGAGTGGTGGTACCCCGAGGGGCGCATG 1320	
QY	1885	GGGGCTTATGTCTGGAGTGTTCAGAGGTTGAGCGCCCGCGGTGGTGGCTGCTTATAGC 1944	
DB	1321	GGCGCTTATGTCTGTAATGTTCAGAGGTTGGGCGCCATGTGTAGCAGCTTACAGC 1380	
QY	1945	TTGGTGTGGGCGAGCGGGGACCTCAAAACCGGGCCCGACACCTGTGTGGGGCTGA 2004	
DB	1381	TTGGTGTGGGCGAGCGAGATGCTCCGAGCCCGGGCCCGACAC --- AGTGGGGGCGGGA 1437	
QY	2005	TTGGTGTGGCTTCTCTGGGTGTCTTGGAGCATCCCTCAGCTCTCTCTGATTGTGCTGC 2064	
DB	1438	CTGGCTGCTCTCTTGGGATTCGCGAGCATCCCTGACTCTCATCTGATTGGTCTGG 1497	
QY	2065	CGTACAGCGTGGCGAGCAGAGGAGCTTCTAGCTAGACAAAGTGGGGCTTAGATCTG 2124	
DB	1498	CGTACAGCGCAGCGCAGAGGGAAGTCTTGGCTAGACAAAGTGGGGCTGGACCTG 1557	
QY	2125	GGGGCTCCACCTTCTGGGACCAAGCTATAGTACGAGCCCTCCCTCTCTCTCGCTGAA 2184	

Db 267 CGTTCCAGTCCCTCACATACAAATTAATCTGTTCTCTCTGTTGATCTGCTCCCTCCAC 326

Qy 241 ACACCTTTACGTGGTCCGACGGATAGCATCTTGGCTTTTAACTCCCTCTCTCTGGGAA 300

Db 327 ACACCTTTATGTTGGCGCGGACACCATCTTGCCTTTATCTCCCTCTCTCAGGGAG 386

Qy 301 AGACCCGGAGGATCGACTGGATGGTACCTGAGACTCACAGACAGAACTCCAGAGAAA 360

Db 387 AGACCCGGCANGATTGACTGGATGGTTCCTGANGCTCACANAGAACTGTAGAGAAA 446

Qy 361 GCGAAGAACAGACGAATCTCAATTTTATCCAGATCTCGCCATTGTCAATGCTCT 420

Db 447 GCGAAGAACAGACGAATCTCAATTTTGTCCAGATCTCGCCATTGCGCAATGCTCT 506

Qy 421 CACCTCTCAGTCGCGACCTTCGCTTTTGTATCCG-AAGTGGGGGTATTGATGTGC 479

Db 507 CACCTCTCAGTGTGACCTTCGCTTTTGTATCCGAAATGCGGGTATTGATGTGC 566

Qy 480 CAGTTCCACAGGTTGAAGACTTGAGACGGCGGGGAAATGCTCTTTTGAGCCAGC 539

Db 567 CAGGTTCCCGNGGTTGAAAACCTTGAATATGCGGGGGGAAATGCTCTTTGAACCACT 626

Qy 540 TCAAGGTCAGAGCTGTAATGCTGGGGCGTCTCTA 578

Db 627 CNGGTCANCACTGTTATTGGCTGGGGGGTCTNTA 665

RESULT 14

AAD08048

ID AAD08048 standard; cDNA; 3293 BP.

XX AC AAD08048;

XX DT

XX TX 06-AUG-2001 (first entry)

XX DE Human extracellular matrix and cell adhesion molecule-4 (XMAD-4) cDNA.

XX KW Human; extracellular matrix and cell adhesion molecule; XMAD;

XX KW gene therapy; genetic disorder; adrenoleukodystrophy; leukaemia;

XX KW Down's syndrome; cystic fibrosis; Gaucher's disease; myotonic dystrophy;

XX KW sickle cell anaemia; thalassaemia; autoimmune disorder; adenocarcinoma;

XX KW inflammatory disorder; acquired immune deficiency syndrome; AIDS;

XX KW Addison's disease; allergy; anaemia; asthma; atherosclerosis; melanoma;

XX KW Crohn's disease; diabetes mellitus; atopic dermatitis; lymphoma; cancer;

XX KW glomerulonephritis; multiple sclerosis; Grave's disease; osteoarthritis;

XX KW osteoporosis; psoriasis; rheumatoid arthritis; ulcerative colitis;

XX KW infection; cell proliferative disorder; actinic keratosis; myeloma;

XX KW arteriosclerosis; neutropenic; anticonvulsant; antithyroid; nephrotropic;

XX KW neuroprotective; dermatological; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT CDS 134..2635

XX FT /*tag= a

XX FT /product= "Human extracellular matrix and cell adhesion

XX FT molecule (XMAD)"

XX FT sig_peptide 134..202

XX FT /*tag= b

XX FT mat_peptide 203..2632

XX FT /*tag= c

XX FT /product= "Mature human extracellular matrix and cell

XX FT adhesion molecule (XMAD)"

XX PN W0200142285-A2.

XX PN 14-JUN-2001.

XX PD

XX XX 05-DEC-2000; 2000WO-US32990.

XX PF

XX PF 10-DEC-1999; 99US-0172852.

XX PR 16-DEC-1999; 99US-0172354.

XX XX

(JNCY-) INCVTE GENOMICS INC.

Yue H, Tang YT, Lal P, Burford N, Azimzai Y, Patterson C;

Baughn MR, Lu DAM, Shah P, Au-Young J;

WPI; 2001-381632/40.

P-PSDB; AAE03640.

New human extracellular matrix and cell adhesion molecules and polynucleotide sequences encoding them, useful for diagnosis, prevention, treatment of genetic, autoimmune and cell proliferative disorders

Claim 5; Page 121-122; 135pp; English.

The present cDNA sequence encodes human extracellular matrix and cell adhesion molecule (XMAD). The XMAD is used for screening a compound for effectiveness as an agonist or antagonist of XMAD. The identified agonist or antagonist are used for treating a disease or condition associated with decreased or increased expression of functional XMAD. The polynucleotides encoding XMAD are useful in somatic or germline gene therapy to correct a genetic deficiency, to express a conditionally lethal gene product and to express a protein which affords protection against intracellular parasites and also for diagnosis of disorders associated with expression of XMAD. They are also used for generating hybridisation probes useful in mapping the naturally occurring genomic sequences and to create knock in humanised animals (pigs) or transgenic animals (mice or rats) to model human diseases. Oligonucleotide or longer fragments derived from the polynucleotide sequences may be used as elements on a microarray. Antibodies which specifically bind XMAD may be used for the diagnosis of disorders associated with the expression of XMAD, or in assays to monitor patients being treated with XMAD. Diseases diagnosed, prevented or treated include genetic disorders such as adrenoleukodystrophy, Down's syndrome, cystic fibrosis, Gaucher's disease, myotonic dystrophy, sickle cell anaemia, thalassaemia, autoimmune/inflammatory disorders such as acquired immune deficiency syndrome (AIDS), Addison's disease, allergies, anaemia, asthma, atherosclerosis, Crohn's disease, diabetes mellitus, atopic dermatitis, glomerulonephritis, multiple sclerosis, Grave's disease, osteoarthritis, osteoporosis, psoriasis, rheumatoid arthritis, ulcerative colitis, bacterial, fungal, parasitic, protozoal and helminthic infections and cell proliferative disorders such as actinic keratosis, arteriosclerosis and cancer including breast, bladder, bone marrow, brain and uterus cancer, leukaemia, adenocarcinoma, lymphoma, melanoma and myeloma.

XX SQ Sequence 3293 BP; 599 A; 978 C; 1052 G; 664 T; 0 other;

Query Match 10.8%; Score 251.8; DB 22; Length 3293;

Best Local Similarity 52.3%; Pred. No. 2.3e-54;

Matches 706; Conservative 0; Mismatches 627; Indels 18; Gaps 6;

Qy 308 GAAGGATCGACTGGATGGTACCTGAGACTCACAGACAGAACTCGAGGAAGAAAGGCAAG- 366

Db 387 GAGCGATCTCTGGGAGGCCCGCTGGAGAGAGAGACTGATGTATCCAGAAAGGGAAGA 446

Qy 367 --AAAGAGGACGAATGTCAAAATTTATCCAGATTTCTGCCATTTGTCAATGCCCTCTCAC 424

Db 447 ACAACACAGACCGAGTGTCTTCAACTTCTCCGCTTCTCCAGAGCCCTTCCACCTCCAC 506

Qy 425 TCCTCAGTGGCGACCTTTCGCTTTTGTATCCGAGTGGGGGTATTGATGTGTCCAGTT 484

Db 507 TGTACGCTGTGGCACCCTACGCTTCCAGGCCCAAGTGCACCTAGTCAACATGCTCACCT 566

Qy 485 TCCA---GCAGTTGAAAGACTTTCAGAGCGCGCGGGGAAATGCTCTTTTGGAGCAGCTC 541

Db 567 TCACCTTTGGAGATGGAGAGATTGACAGATGGGAGGCAAGTGTCCCTATGACCCAGCTA 626

Qy 542 AACCGTCAGCAGCTGTAAATGGCTGGGGCGCTCTCTTACACCGGCACCTGTGAAGAACTCC 601

Db 627 AGGSCCATGCTGGCCTTCTTTGTGATGGTGAAGCTGTACTCGGCCACACTCAACAACCTCC 686

Qy 602 TGGGGACTGAGCCCATCATCTCCGAGCTGTGGGTCCGAGTGGAGTGGATTGCAACAG 661

Db 687 TGGGACGGAACCCATTATCTCGCTAACATGGGCG---CCGACCACTCCATGAAGACAG 743
QY 652 AGACCTTGTCATCTCGGCTTAATGCTCCAGCCTTTGTGCGAGCTATGCTCTGTGAGCCGAG 721
Db 744 AGTACCTGGCCTTTTGGCTCAAGGAACCTCACTTTGTAGGCTCTGCCTATGTACCTGAGA 803
QY 722 CTGAGTGGGGGATGAAGATGAGAGCATGAATCTTTTTCCTTTCACGGAGACCTCCC 781
Db 804 CTGTGGGAGCTTTCACGGGGACGACGACAAGGTCTACTTCTTTCACGGAGCGGGCAC 863
QY 782 GAGTCTTGACTCTCTATGAGCGCATCAAGGTCCCAAGAGTGGCCGAGCTGTGTGGGGG 841
Db 864 TGGAGTCCGACTGCTATCCGAGCAGGTGGTGTGTGGCCGCTGTCTGCAAGGGCG 923
QY 842 ACCTTGGGGGAGGAGACCTTTCAGCAGAGATGGACGACGTTTCTGAAGGCTGACCTGC 901
Db 924 ATATGGGGGGCGCAGGACCTTGCAGAGGAAGTGGACACGTTCTCTGAAGGCGCGCTGG 983
QY 902 TGTGCCAGGGCCGAGCATGGCCGGGCTCGGGGGTTCGTGAGGCTATGCGAGAGTTC 961
Db 984 CATGCTCTGCCCGGAACCTGGCAGCTCTACTTCAACCACTGCAGGCGATG---CACACC 1040
QY 962 GGCCTCAGCTTGGAGCGGAACCCCATCTTTTATGGGATCTTTCTCCAGTGGGAAG 1021
Db 1041 TGCAGGACCTCTCGGCACACACACCCTTCTTTGGGGTTTTCAAGCACAGTGGGTG 1100
QY 1022 GAGCTGCCATCTCTGTGTGTGCTTCCGACCCCAAGACATCCGGGCGAGTGTGAATG 1081
Db 1101 ACATGTACTCTGCGGCCATCTGTGAGTACCAGTTTGAAGAGATCCAGCGGTGTGAGG 1160
QY 1082 GTCCCTTTAGAGAGTAAACATGACTGCACAGGGGACTGCGTGTCAATGACACAGG 1141
Db 1161 GCCCTATAGAGTACCATGAGGAAGCCCAAGTGGGACCGCTACACACTGACCCCT---G 1217
QY 1142 TGCCCCAGCCAGACTGAGAGTGCATCGGCACACATGAAGCTCCAGCAGTTGGAT 1201
Db 1218 TACCAGCCCTCGGCTGGCTGCTGCTATTAACAAGTGGCATCGGGCCACCGCTACACCA 1277
QY 1202 COTCACTCTCCCTGCCAGCGGCTGCTACCTTTATCAGAGACCAACCTCTCATGGACA 1261
Db 1278 GCTCCCTGGAGTACCCGACAACTCCTCAACTTGTCTCAAGAAGCACCCGCTGATGGAG 1337
QY 1262 GCGCGCTGTTCCCGCTCAGCGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1321
Db 1338 AGCAGTGGGCGCTCGGTGGAGCGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1397
QY 1322 GAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1381
Db 1398 ACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1457
QY 1382 GGACAGAGATGGACACCTCCACCGGGCTGTGCGCATTTGGAGCTCAGCTCAGTGTCTTGG 1441
Db 1458 GCACAGGAGCGGCTGCTGCTCAAGGCTGTGAGCGCTGGGCGCTTCCACCTGATTG 1517
QY 1442 AGGATCTGGCTTGT---CCGAGACACAGCGCGTGTGAGCAGTGAATTTGACACAG 1498
Db 1518 AGGAGCTCAGCTGTTTACAGGAGCGCCATGAGAAAGCCTGTGCTATCTCAGAGCAAGA 1577
QY 1499 ATTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1558
Db 1578 AGCTGCTCTTGGCGGCTCCGCTCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1637
QY 1559 GTCCTCAGAGTGTCTGAGGTGTATCTTGGCCCGAGACCCGCTGTGCGCCTGGAGCTTCC 1618
Db 1638 AGTATCGCTCTGTGACAGCTGTGCTTCCCTGCGCGGAGCCCTTATTGCGCCTGGAGCGTCA 1697
QY 1619 GCCTTGATGCTGTGTGCGCCACGCGCGGA 1649
Db 1698 ACACAGCGCTGTGTGCGCGTGGGTGCCA 1728

RESULT 15
AAC84887

AC84887 standard; cDNA; 2155 BP.
AAC84887;
20-APR-2001 (first entry)
Human SEC6 nucleic acid sequence (clone ID 20422974.0.132).
SEC6; cytostatic; gynecological; gene therapy; screening assay; human;
chromosomal mapping; forensic biology; cell proliferation; cancer;
cell differentiation; immune associated disorder; gestational disease;
SEC6; ss.
Homo sapiens.
Key Location/Qualifiers
CDS 166..1937
/*tag= a
/product= "SEC6"
WO200078802-A2.
28-DEC-2000.
23-JUN-2000; 2000WO-US17328.
23-JUN-1999; 99US-0140584.
20-JUL-1999; 99US-0144722.
16-SEP-1999; 99US-0154520.
22-JUN-2000; 2000US-0604286.
(CURA-) CURAGEN CORP.
Shimkets RA, Fernandes E, Vernet C, Yang M, Boldog FL;
Herrmann JL;
WPI; 2001-071385/08.
P-PSDB; AAB48373.
Polynucleotides encoding SECX proteins useful for treating disease
characterized by an aberrant level of cell proliferation and/or
differentiation like cancer or immune associated disorders -
Claim 3; Fig 6; 132pp; English.
The invention relates to human SECX polypeptides and polynucleotides
encoding them. The SECX polypeptides can be expressed by standard
recombinant methodology. The SECX polypeptides are useful for treating
or preventing a SECX-associated disorder. The invention is useful in
screening assays; detection assays (e.g. chromosomal mapping, cell and
tissue typing, forensic biology); predictive medicine (diagnostic assays,
prognostic assays, monitoring clinical trials, and pharmacogenomics); and
methods of treatment (e.g. therapeutic and prophylactic), especially
disorders characterized by aberrant cell proliferation and/or
differentiation like cancer or immune associated disorders or gestational
disease. The present sequence represents a SEC6 nucleic acid sequence.
Sequence 2155 BP; 444 A; 614 C; 645 G; 452 T; 0 other;
Query Match 10.7%; Score 248.6; DB 22; Length 2155;
Best Local Similarity 52.1%; Pred. No. 1.3e-53;
Matches 704; Conservative 0; Mismatches 629; Indels 18; Gaps 6;
QY 308 GAAGATGCTGATGGTACCTGAGACTCAGACAGAACTGCGAGGAAGCAAG- 366
Db 419 GAGCGATCTCTGGGAGGCCCGCTGGAGAAAGAGACTGAGTGTATCCAGAAAGGAAGA 478
QY 367 --AAAGAGGAGAAATGTACAAATTTTATCCAGATTTCTGCCCATTTGCTCAATGCTCTCACC 424
Db 479 ACAACCAAGACGAGTGTCTTCAACTTCATCCGCTTCTCTGACGCCCTACAAATGCTCCACC 538
QY 425 TCCTCAGTGGCGGACCTTCGCTTTTGATCCGAACTGCGGGTATTATGATGTGTCCAGTT 484

Db	539	TGTAGTCTGTGGACACTACGCTTCCAGCCCAAGTGCAOCTACGTCAACATGTCACCT	598
Qy	485	TCCA---GCAGGTTGAAGACTTGAGAGCGCCGGGGAAATGTCCTTTTTCAGCCAGCTC	541
Db	599	TCACCTTTGGACATGGAGAGTTGAAGATGGAGGGCAAGTGTCCTATACCCAGCTA	658
Qy	542	AACGGTCAGACGTGTAATGGCTGGGGCGTCTCTACACCGCCACACTGTGAAGAACTTCC	601
Db	659	AGGCCCATGCTGGGCTTCTTTGTGATGGTGAAGTGTACTCGGCCACACATCAACAACTTCC	718
Qy	602	TGGGACTTGACCCATCATCTCCCGAGCTGTGGCTCGAGCTGAGGACTGGATTGCAACAG	661
Db	719	TGGGACGGAACCATATCTCTGCTAACATCTGGGC---CCACCACTCCATGAAGACAG	775
Qy	662	AGACCTTTGTCACTCTGGCTTAATGCTCCAGCCTTTTGTGCGACGTATGGCTCTGAGCCCGAG	721
Db	776	AGTACCTGGCCTTTTGGCTCAACGAACCTCACTTTGTAGGCTCTGCCCTATCTACCTGAGA	835
Qy	722	CTGAGTGGGGGATGAAGATGGAGACGATGAAATCTTTTTCCTTTCACGGAGACCTCCC	781
Db	836	GTGTGGGAGCTTCACGGGGGACGACACAAGGTCCTACTTCTTTCAGGGAGGGGCGAG	895
Qy	782	GAGTGTGGGACTCTATGAGCGCATCAAGTCCCAAGAGTGGCCGAGTGTGCGGGGG	841
Db	896	TGGAGTCGCACTGCTATGTCGAGCAGGTGGTGGCTCGTGTGGCCCGTGTCTGCAAGGGCG	955
Qy	842	ACCTTTGGGGGAGGAAGACCCCTTCAGCAGAGATGGAGCAGCTTTCTGAAGCTGACCTGC	901
Db	956	ATATGGGGGGCGCAGGACCTTCAGAGGAAGTGGACCACGTTCTGAAAGCGCGGTGG	1015
Qy	902	TGTGCCAGGGCCCGAGCATGGCCGGGCTCCGGGGTTCGACGGCTATGCGAGGCTTC	961
Db	1016	CATGCTCTGCCCGCAACTGGCAGCTCTACTTCAACACAGCTCAGSGGATG---CACACCC	1072
Qy	962	GGCCTCAGCCTGGAGCGGGAACCCCATCTTTTATGGATCTTTTCTCCAGTGGGAAG	1021
Db	1073	TGCAGGACACTCTCTGGCACAAACACCACTTCTTTTGGGGTTTTCAGCACAGTGGGGTG	1132
Qy	1022	GAGCTGCCATCTCTGTGTGTGCTTCCGACCCCAAGACATCCGGGCGAGTGTGAATG	1081
Db	1133	ACATGTACCTGTGCGCCATCTGTGATACCACTTGGAGAGATCCAGGGGTGTTTGAGG	1192
Qy	1082	GTCCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCCCTGTATGGAACAACGAGG	1141
Db	1193	GCCCTATAGAGAGTACCATGAGGAAGCCAGAAAGTGGGACCGGTACACTGACCT---G	1249
Qy	1142	TGCCCCAGCCCGAGCTGGAGAGTGCATCGGCCCAACACATGAAGCTCCAGCAGTTGGAT	1201
Db	1250	TACCCAGCCCTCGGCTGGCTCGTGCATTAACAACCTGGCATCGGGCGCCACGGCTACACCA	1309
Qy	1202	CCTCACTCTCCCTGCCAGACCGCGTGTCACTTTTATCAGAGACACCCCTCTCATGGACA	1261
Db	1310	GCTCCCTGGAGCTACCCGACAACATCTCACTTCGTCAAGAAGCACCCGCTGATGGAGG	1369
Qy	1262	GGCCCGTGTCCCGGCTGACCGCCGCCCTGCTGGTGCATCAGATACAGCCCTATCTCA	1321
Db	1370	AGCAGTGGGGCTCGGTGGAGCGCGCCCTGCTGCTGAAGAAGGGCACCAACTTCACCC	1429
Qy	1322	GAGTGTGGGGCCAGGGTGACAGCTCTCAGGGAAGATATGACGTGCTCTACCTGG	1381
Db	1430	ACCTGGTGGCGAGCGGGTTACAGGACTGTGAGGACCCTATACAGTGTCTGTCAATTG	1489
Qy	1382	GGACAGAGATGGACATCCACCGGGCTGTGCGCATTTGGAGCTCAGCTCAGTGTCTTGG	1441
Db	1490	GCACAGGAGACGGCTGGCTCAAGGCTGTGAGCCTGGGGCCCTGGGTTTCACTGATTG	1549
Qy	1442	AGGATCTGGCCTTGT---CCCAAGAACACACCGGTTGAGAGCATGAAATTTGTACCACG	1498
Db	1550	AGGAGCTCAGCTGTTTGACCAGAGGCCCATGAGAAAGCTGGTCTCTCAGAGCAAGA	1609
Qy	1499	ATTGGCTCTCTGGTGGGCTCCCATACTGAGGTGACACAAGTGAACACAGCAACTGTGGCC	1558
Db	1610	AGCTGCTCTTTTGGCGGCTCCCGCTCTCAGCTGGTGGAGCTGCCGCTGGCCGACTGCATPAA	1669

Qy 1559 GTCTCCAGAGCTGCTCGGAGTGTATCTTGGCCAGGACCCGCTGTGCGCCTGGAGCTTCC 1618
Db 1670 AGTATCGCTCCTGTGCAGACTGTGTCTCGCCCGGAGCCCTATTGCGCCTGGAGCGTCA 1729
Qy 1619 GGCTTGATGCTTGTGTGCCCCACGCCGGGA 1649
Db 1730 ACACCAGCCGCTGTGTGGCCGTGGGTGGCCA 1760

Search completed: September 25, 2003, 19:53:08
Job time : 575.81 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2003, 19:27:00 ; Search time 6789.65 Seconds
(without alignments)
11385.680 Million cell updates/sec

Title: US-09-284-180A-2
Perfect score: 2331
Sequence: 1 atcttgcagggccgagcg.....gtgatgagacctcatctaa 2331

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 33363688 seqs, 16581889874 residues

Total number of hits satisfying chosen parameters: 66727376

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main:*

1: /cgn2_6/ptodata/2/pna/PCRTUS_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/PCRTUS_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US06_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US07_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US080_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US081_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US083_COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US084_COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US085_COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US086_COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US088_COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US089_COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US090_COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US091_COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US092A_COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US092B_COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US093A_COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US093B_COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US094_COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US095A_COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US095B_COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US095C_COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US095D_COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US096A_COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US096B_COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US096C_COMB.seq.*
29: /cgn2_6/ptodata/2/pna/US096D_COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US096E_COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US097A_COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US097B_COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US097C_COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US098A_COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US098B_COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US098C_COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US098D_COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US099A_COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US099B_COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US099C_COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US099D_COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US099E_COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US099F_COMB.seq.*

44: /cgn2_6/ptodata/2/pna/US100A_COMB.seq.*
45: /cgn2_6/ptodata/2/pna/US100B_COMB.seq.*
46: /cgn2_6/ptodata/2/pna/US101A_COMB.seq.*
47: /cgn2_6/ptodata/2/pna/US101B_COMB.seq.*
48: /cgn2_6/ptodata/2/pna/US102A_COMB.seq.*
49: /cgn2_6/ptodata/2/pna/US102B_COMB.seq.*
50: /cgn2_6/ptodata/2/pna/US103A_COMB.seq.*
51: /cgn2_6/ptodata/2/pna/US103B_COMB.seq.*
52: /cgn2_6/ptodata/2/pna/US104A_COMB.seq.*
53: /cgn2_6/ptodata/2/pna/US104B_COMB.seq.*
54: /cgn2_6/ptodata/2/pna/US6000_COMB.seq.*
55: /cgn2_6/ptodata/2/pna/US6001_COMB.seq.*
56: /cgn2_6/ptodata/2/pna/US6002_COMB.seq.*
57: /cgn2_6/ptodata/2/pna/US6003_COMB.seq.*
58: /cgn2_6/ptodata/2/pna/US6004_COMB.seq.*
59: /cgn2_6/ptodata/2/pna/US6005_COMB.seq.*
60: /cgn2_6/ptodata/2/pna/US6006_COMB.seq.*
61: /cgn2_6/ptodata/2/pna/US6007_COMB.seq.*
62: /cgn2_6/ptodata/2/pna/US6008_COMB.seq.*
63: /cgn2_6/ptodata/2/pna/US6009_COMB.seq.*
64: /cgn2_6/ptodata/2/pna/US6010_COMB.seq.*
65: /cgn2_6/ptodata/2/pna/US6011_COMB.seq.*
66: /cgn2_6/ptodata/2/pna/US6012_COMB.seq.*
67: /cgn2_6/ptodata/2/pna/US6013_COMB.seq.*
68: /cgn2_6/ptodata/2/pna/US6014_COMB.seq.*
69: /cgn2_6/ptodata/2/pna/US6015_COMB.seq.*
70: /cgn2_6/ptodata/2/pna/US6016_COMB.seq.*
71: /cgn2_6/ptodata/2/pna/US6017_COMB.seq.*
72: /cgn2_6/ptodata/2/pna/US6018_COMB.seq.*
73: /cgn2_6/ptodata/2/pna/US6019_COMB.seq.*
74: /cgn2_6/ptodata/2/pna/US6020_COMB.seq.*
75: /cgn2_6/ptodata/2/pna/US6021_COMB.seq.*
76: /cgn2_6/ptodata/2/pna/US6022_COMB.seq.*
77: /cgn2_6/ptodata/2/pna/US6023A_COMB.seq.*
78: /cgn2_6/ptodata/2/pna/US6023B_COMB.seq.*
79: /cgn2_6/ptodata/2/pna/US6024_COMB.seq.*
80: /cgn2_6/ptodata/2/pna/US6025_COMB.seq.*
81: /cgn2_6/ptodata/2/pna/US6026_COMB.seq.*
82: /cgn2_6/ptodata/2/pna/US6027_COMB.seq.*
83: /cgn2_6/ptodata/2/pna/US6028_COMB.seq.*
84: /cgn2_6/ptodata/2/pna/US6029_COMB.seq.*
85: /cgn2_6/ptodata/2/pna/US6030_COMB.seq.*
86: /cgn2_6/ptodata/2/pna/US6031_COMB.seq.*
87: /cgn2_6/ptodata/2/pna/US6032_COMB.seq.*
88: /cgn2_6/ptodata/2/pna/US6033_COMB.seq.*
89: /cgn2_6/ptodata/2/pna/US6034_COMB.seq.*
90: /cgn2_6/ptodata/2/pna/US6035_COMB.seq.*
91: /cgn2_6/ptodata/2/pna/US6036_COMB.seq.*
92: /cgn2_6/ptodata/2/pna/US6037_COMB.seq.*
93: /cgn2_6/ptodata/2/pna/US6038_COMB.seq.*
94: /cgn2_6/ptodata/2/pna/US6039_COMB.seq.*
95: /cgn2_6/ptodata/2/pna/US6040_COMB.seq.*
96: /cgn2_6/ptodata/2/pna/US6041_COMB.seq.*
97: /cgn2_6/ptodata/2/pna/US6042_COMB.seq.*
98: /cgn2_6/ptodata/2/pna/US6043_COMB.seq.*
99: /cgn2_6/ptodata/2/pna/US6044_COMB.seq.*
100: /cgn2_6/ptodata/2/pna/US6045_COMB.seq.*
101: /cgn2_6/ptodata/2/pna/US6046_COMB.seq.*
102: /cgn2_6/ptodata/2/pna/US6047_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2331	100.0	2331	18	US-09-284-180-2
2	2331	100.0	2331	18	US-09-284-180A-2
3	2331	100.0	4008	18	US-09-284-180-1
4	2331	100.0	4008	18	US-09-284-180A-1

Sequence 2, Appli
Sequence 2, Appli
Sequence 1, Appli
Sequence 1, Appli

[illegible]

QY 541 CAACGGTCACAGCTGAATGGCTGGGGCGTCTCTACACGGCCACTGTGAAGAACTTC 600
DB |||||||
QY 541 CAACGGTCACAGCTGAATGGCTGGGGCGTCTCTACACGGCCACTGTGAAGAACTTC 600
DB |||||||
QY 601 CTGGGGACTGAGCCATCATCTCCGAGCTGTGGTCGAGCTGAGGACTGATTCGAACTTC 660
DB |||||||
QY 601 CTGGGGACTGAGCCATCATCTCCGAGCTGTGGTCGAGCTGAGGACTGATTCGAACTTC 660
DB |||||||
QY 661 GAGACCTGTTCATCTGGCTTAATGTCTCCAGCTTTTGTGCGAGCTATGTCCTGAGCCCA 720
DB |||||||
QY 661 GAGACCTGTTCATCTGGCTTAATGTCTCCAGCTTTTGTGCGAGCTATGTCCTGAGCCCA 720
DB |||||||
QY 721 GCTGAGTGGGGGATGAAGATGAGAGCATGAATCTTTTTCACGGAGACTCTC 780
DB |||||||
QY 721 GCTGAGTGGGGGATGAAGATGAGAGCATGAATCTTTTTCACGGAGACTCTC 780
DB |||||||
QY 781 CGAGTGTGGACTCTCATGAGCGCATCAAGGTCCCAAGAGTGGCCGAGTGTCTGGGG 840
DB |||||||
QY 781 CGAGTGTGGACTCTCATGAGCGCATCAAGGTCCCAAGAGTGGCCGAGTGTCTGGGG 840
DB |||||||
QY 841 GACCTGGGGGAGGAGACCTTTCAGCAGAGATGGACGACGTTTCTGAAGGCTGACCTG 900
DB |||||||
QY 841 GACCTGGGGGAGGAGACCTTTCAGCAGAGATGGACGACGTTTCTGAAGGCTGACCTG 900
DB |||||||
QY 901 CTGTGCCCCAGGGCCGAGCATGCGCGGCTCCGGGTTCTGAGGCTATGGCAGAGCTT 960
DB |||||||
QY 901 CTGTGCCCCAGGGCCGAGCATGCGCGGCTCCGGGTTCTGAGGCTATGGCAGAGCTT 960
DB |||||||
QY 961 CGGCTCAGCTCGAGCGGGAACCCCACTTTTATGGGATCTTTCTCCCACTGGAA 1020
DB |||||||
QY 961 CGGCTCAGCTCGAGCGGGAACCCCACTTTTATGGGATCTTTCTCCCACTGGAA 1020
DB |||||||
QY 1021 GGAGTGCATCTCTGCTGTGTGCTTCCGACCCCAAGACATCCGGGCACTGCTGAAT 1080
DB |||||||
QY 1021 GGAGTGCATCTCTGCTGTGTGCTTCCGACCCCAAGACATCCGGGCACTGCTGAAT 1080
DB |||||||
QY 1081 GGTCCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCCTGTCTATGACACAGAG 1140
DB |||||||
QY 1081 GGTCCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCCTGTCTATGACACAGAG 1140
DB |||||||
QY 1141 GTGCCCCAGCCACAGCTGGAGATGCATCGCCCAACAAATGAAGCTCCAGCAGTTTGA 1200
DB |||||||
QY 1141 GTGCCCCAGCCACAGCTGGAGATGCATCGCCCAACAAATGAAGCTCCAGCAGTTTGA 1200
DB |||||||
QY 1201 TCTCACTCTCCCTGCGACAGCGGTGCTCACTTTATCAGACACACCTCTCATGGAC 1260
DB |||||||
QY 1201 TCTCACTCTCCCTGCGACAGCGGTGCTCACTTTATCAGACACACCTCTCATGGAC 1260
DB |||||||
QY 1261 AGGCCGCTGTCCCGCTCAGCGCCGCCCTGCTGGTCACTACAGATACAGCCTATCTC 1320
DB |||||||
QY 1261 AGGCCGCTGTCCCGCTCAGCGCCGCCCTGCTGGTCACTACAGATACAGCCTATCTC 1320
DB |||||||
QY 1321 AGAGTCGTGGCCACAGGTGACAGCTCTCAGGGAAAGATATGACGTCTCTACCTG 1380
DB |||||||
QY 1321 AGAGTCGTGGCCACAGGTGACAGCTCTCAGGGAAAGATATGACGTCTCTACCTG 1380
DB |||||||
QY 1381 GGGACAGAGATGGACACCTCCACCGGCTGTGCGCATTTGGAGCTCAGTCTGCTTG 1440
DB |||||||
QY 1381 GGGACAGAGATGGACACCTCCACCGGCTGTGCGCATTTGGAGCTCAGTCTGCTTG 1440
DB |||||||
QY 1441 GAGGATCTGGCTTGTTCAGAACACAGCGGTTGAGAGCATGAATTTGACACAGAT 1500
DB |||||||
QY 1441 GAGGATCTGGCTTGTTCAGAACACAGCGGTTGAGAGCATGAATTTGACACAGAT 1500
DB |||||||
QY 1501 TGGCTCTGGTGGGCTCCCATACTAGGTGACACAAAGTGAACACCAAGCACTGTGGCGCT 1560
DB |||||||
QY 1501 TGGCTCTGGTGGGCTCCCATACTAGGTGACACAAAGTGAACACCAAGCACTGTGGCGCT 1560
DB |||||||
QY 1561 CTCCAGAGCTGCTCGAGTGTATCTGCCCCAGGACCCCGTGTGGCCTGGAGCTTCGG 1620
DB |||||||
QY 1561 CTCCAGAGCTGCTCGAGTGTATCTGCCCCAGGACCCCGTGTGGCCTGGAGCTTCGG 1620
DB |||||||

QY 1621 CTTGATGCTTGTGTGGCCACGCCGGCGAGCACCGGGATGGTTCAAGATATAGAGTCA 1680
DB |||||||
QY 1621 CTTGATGCTTGTGTGGCCACGCCGGCGAGCACCGGGATGGTTCAAGATATAGAGTCA 1680
DB |||||||
QY 1681 GCGGATGCTCTCTTCTTGTGTCCAAAGAACCTTGGAGAACATCCCGTAGTGTGAAGTT 1740
DB |||||||
QY 1681 GCGGATGCTCTCTTCTTGTGTCCAAAGAACCTTGGAGAACATCCCGTAGTGTGAAGTT 1740
DB |||||||
QY 1741 CCGGTGGCTACTGTGGGCCACGTGGTCTGCCATGTTCCCCAGTGTCTGCTGGCATCC 1800
DB |||||||
QY 1741 CCGGTGGCTACTGTGGGCCACGTGGTCTGCCATGTTCCCCAGTGTCTGCTGGCATCC 1800
DB |||||||
QY 1801 TGTGTGTGCCACAGCCAGTGGAGTACTGCGCTCACTCCCGGAGGATGGACTAGAG 1860
DB |||||||
QY 1801 TGTGTGTGCCACAGCCAGTGGAGTACTGCGCTCACTCCCGGAGGATGGACTAGAG 1860
DB |||||||
QY 1861 GTGGTGGTGACCCACAGGGCCATGGGGCTTATGCTTGGAGTGTTCAGGAGGTGGAGCC 1920
DB |||||||
QY 1861 GTGGTGGTGACCCACAGGGCCATGGGGCTTATGCTTGGAGTGTTCAGGAGGTGGAGCC 1920
DB |||||||
QY 1921 GCCCGCTGTGGCTCTTATAGCTTGGTGGGCGAGCCAGCGGGACCCCTCAAAACCGG 1980
DB |||||||
QY 1921 GCCCGCTGTGGCTCTTATAGCTTGGTGGGCGAGCCAGCGGGACCCCTCAAAACCGG 1980
DB |||||||
QY 1981 GCCCACACCGTGTGGGGCTGGATTGGTGGCTTCTCTGGGTGTCTTTCAGCATCC 2040
DB |||||||
QY 1981 GCCCACACCGTGTGGGGCTGGATTGGTGGCTTCTCTGGGTGTCTTTCAGCATCC 2040
DB |||||||
QY 2041 CTCACCTCTCTCTGATTTGGTCCCGTCAACAGCTTCTGGGACCAAGCTATAGTCAG 2100
DB |||||||
QY 2041 CTCACCTCTCTCTGATTTGGTCCCGTCAACAGCTTCTGGGACCAAGCTATAGTCAG 2100
DB |||||||
QY 2101 AGAGACAAAGTGGGCTTAGATCTGGGGCTTCCACCTTCTGGGACCAAGCTATAGTCAG 2160
DB |||||||
QY 2101 AGAGACAAAGTGGGCTTAGATCTGGGGCTTCCACCTTCTGGGACCAAGCTATAGTCAG 2160
DB |||||||
QY 2161 GACCTCTCTCTCTTGGCTGAAGATGAACAGCTGCCCTGGCTGGGTAAAGCGGGC 2220
DB |||||||
QY 2161 GACCTCTCTCTCTTGGCTGAAGATGAACAGCTGCCCTGGCTGGGTAAAGCGGGC 2220
DB |||||||
QY 2221 AGTGGTTTGGTGGCTTCCCTCCACCTTCTGCTGATTTGCTTGGCCAGCCAGCCAC 2280
DB |||||||
QY 2221 AGTGGTTTGGTGGCTTCCCTCCACCTTCTGCTGATTTGCTTGGCCAGCCAGCCAC 2280
DB |||||||
QY 2281 ATCCGGCTCAGTGGGGCGCTCTAGCCAGCTGTGATGAGACCTTCCATCTAA 2331
DB |||||||
QY 2281 ATCCGGCTCAGTGGGGCGCTCTAGCCAGCTGTGATGAGACCTTCCATCTAA 2331
DB |||||||

RESULT 3

US-09-284-180-1
; Sequence 1, Application US/09284180
; GENERAL INFORMATION:
; APPLICANT: KIMURA, Kaoru
; APPLICANT: KIMURA, Kaoru
; TITLE OF INVENTION: NOVEL SEMAPHORIN GENE: SEMAPHORIN W
; FILE REFERENCE: 20-456P
; CURRENT APPLICATION NUMBER: US/09/284,180
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4008
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(75)
; FEATURE:
; OTHER INFORMATION: Coding region from residue 76 to 2406
; NAME/KEY: 3'UTR
; LOCATION: (1)..(3977)


```

; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: ( )..(4008)
US-09-284-180-1

```

```
Query Match      100.0%; Score 2331; DB 18; Length 4008;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

[illegible]

Qy	961	CGG	CCTCAGCCTTGGAGCGGAACCCCATCTTTTATGGATCTTTTCTCCCACTGGGAA	1021
Db	1036	CGG	CCTCAGCCTTGGAGCGGAACCCCATCTTTTATGGATCTTTTCTCCCACTGGGAA	1095
Qy	1021	GGAG	CTGCCATCTCTGCTGTGTGCGCTTCGACCCCAACACATCCGGGCAGTCTCAAT	1080
Db	1096	GGAG	CTGCCATCTCTGCTGTGTGCGCTTCGACCCCAACACATCCGGGCAGTCTCAAT	1155
Qy	1081	GGT	CCCTTTAGAGAGCTAAACATGACTGTCAACAGGGGACTGCGTGTATCGACACAGGAG	1140
Db	1156	GGT	CCCTTTAGAGAGCTAAACATGACTGTCAACAGGGGACTGCGTGTATCGACACAGGAG	1215
Qy	1141	GTG	CCCCAGCCAGACCTTGGAGAGTGATCGCCAAACAATGAAGTCTCCAGCAGTTTGGGA	1200
Db	1216	GTG	CCCCAGCCAGACCTTGGAGAGTGATCGCCAAACAATGAAGTCTCCAGCAGTTTGGGA	1275
Qy	1201	TCCT	CACCTCCCTGCCAGACCGGCTCACCCTTTATCAGAGACCAACCTCTCATCGGAC	1260
Db	1276	TCCT	CACCTCCCTGCCAGACCGGCTCACCCTTTATCAGAGACCAACCTCTCATCGGAC	1335
Qy	1261	AGG	CCCTGTTCGGCGGTGACGGCGGCCCTCTGCTGTCTACTACAGATACAGGCTATCTC	1320
Db	1336	AGG	CCCTGTTCGGCGGTGACGGCGGCCCTCTGCTGTCTACTACAGATACAGGCTATCTC	1395
Qy	1321	AGAG	CTGTGCCACAGGGTGACAGCCTCTCAGGGAAAGAAATATGACGTGCTTACCTG	1380
Db	1396	AGAG	CTGTGCCACAGGGTGACAGCCTCTCAGGGAAAGAAATATGACGTGCTTACCTG	1455
Qy	1381	GGG	ACAGAGATGGACACCTCCACCGGGCTGTGGGCATTCGAGCTCAGCTCAGTGTCTTG	1440
Db	1456	GGG	ACAGAGATGGACACCTCCACCGGGCTGTGGGCATTCGAGCTCAGCTCAGTGTCTTG	1515
Qy	1441	GAG	GATCTGGCTTGTTCACAGAACACACCGCGTTGAGAGCATGAAATTTGTACCACGAT	1500
Db	1516	GAG	GATCTGGCTTGTTCACAGAACACACCGCGTTGAGAGCATGAAATTTGTACCACGAT	1575
Qy	1501	TGG	CTCTGTGGGCTCCCATCTAGAGGTGACACAAAGTGAACACCACTGTGGCCGT	1560
Db	1576	TGG	CTCTGTGGGCTCCCATCTAGAGGTGACACAAAGTGAACACCACTGTGGCCGT	1635
Qy	1561	CTC	CAGAGCTGTCTGGAGTGTATCTTGGCCACAGGACCCCGTGTGGCCTGGAGGTTC	1620
Db	1636	CTC	CAGAGCTGTCTGGAGTGTATCTTGGCCACAGGACCCCGTGTGGCCTGGAGGTTC	1695
Qy	1621	CTT	GATCTTGTGGGCCACGCGGGGAGCACCGCGGATGTTCAAGATATAGAGTCA	1680
Db	1696	CTT	GATCTTGTGGGCCACGCGGGGAGCACCGCGGATGTTCAAGATATAGAGTCA	1755
Qy	1681	CGG	ATCTCTCTTTTGTCTCAAAAAGAACCTCGAGAACATCCCGTAGTGTGTAAGTT	1740
Db	1756	CGG	ATCTCTCTTTTGTCTCAAAAAGAACCTCGAGAACATCCCGTAGTGTGTAAGTT	1815
Qy	1741	CCG	GTGCTACTGTGGGCCAGTGGTCTGCCATGTTCGCCAGTTCTGCTGGGCATCC	1800
Db	1816	CCG	GTGCTACTGTGGGCCAGTGGTCTGCCATGTTCGCCAGTTCTGCTGGGCATCC	1875
Qy	1801	TGT	GTGGCACAGCCACTGGAGTGACTGCGCTACTCCCGGGAGGATGACATAGAG	1860
Db	1876	TGT	GTGGCACAGCCACTGGAGTGACTGCGCTACTCCCGGGAGGATGACATAGAG	1935
Qy	1861	GTG	GTGGTACCCACGGGGCCATGGGGCTTATGCTTGCAGTCTCAGGAGGTGGAGCC	1920
Db	1936	GTG	GTGGTACCCACGGGGCCATGGGGCTTATGCTTGCAGTCTCAGGAGGTGGAGCC	1995
Qy	1921	GCC	CGGTGTGGCTGCTTATAGCTTGGTGTGGGACCGCAGCGGGACCCCTCAAAACCGG	1980
Db	1996	GCC	CGGTGTGGCTGCTTATAGCTTGGTGTGGGACCGCAGCGGGACCCCTCAAAACCGG	2055
Qy	1981	GCC	CACACCGTGTGGGGCTGATGGTTGGCTTCTCTCTGGGTCTTCTTGAGCATCC	2040
Db	2056	GCC	CACACCGTGTGGGGCTGATGGTTGGCTTCTCTCTGGGTCTTCTTGAGCATCC	2115
Qy	2041	CTC	ACTCTCTCTCTGATTGTCGCGCTCAGCAGCGTCTGGCGACAGAGGGAGTCTTAGCT	2100

QY	1321	AGAGTCGTGGCCACACAGGGTGACCAAGCCTCTCAGGGAAAGAATATGACGTGCTCTACCTG	1380
DB	1396	AGAGTCGTGGCCACACAGGGTGACCAAGCCTCTCAGGGAAAGAATATGACGTGCTCTACCTG	1455
QY	1381	GGACAGAGGATGGACACCTCCACACCGGGCTGTCGCATTTGGAGCTCAGGTCAGTGCTCTG	1440
DB	1456	GGACAGAGGATGGACACCTCCACACCGGGCTGTCGCATTTGGAGCTCAGGTCAGTGCTCTG	1515
QY	1441	GAGGATCTGGCCTTTGTTCCAGAACACAGCGGGTTGAGAGCATGAAAATGTTACCAACAGAT	1500
DB	1516	GAGGATCTGGCCTTTGTTCCAGAACACACAGCGGGTTGAGAGCATGAAAATGTTACCAACAGAT	1575
QY	1501	TGGCTCCTGGTGGGCTCCCATCTAGGTGACACAAGTGAACACCAAGCACTGTGGCCGT	1560
DB	1576	TGGCTCCTGGTGGGCTCCCATCTAGGTGACACAAGTGAACACCAAGCACTGTGGCCGT	1635
QY	1561	CTCCAGAGCTGCTCGGAGTGATCCTTGGCCAGGACCCGCTGTGCGCCTTGAGCTTCCCGG	1620
DB	1636	CTCCAGAGCTGCTCGGAGTGATCCTTGGCCAGGACCCGCTGTGCGCCTTGAGCTTCCCGG	1695
QY	1621	CTTTGATGCTGTGTGGCCCAACGCGGGAGCACCGGGGATGGTTCAAGATATAGAGTCA	1680
DB	1696	CTTTGATGCTGTGTGGCCCAACGCGGGAGCACCGGGGATGGTTCAAGATATAGAGTCA	1755
QY	1681	GCGGATGCTCTCTTTGTTGTCCAAAAGAACCTGGAGAACATCCCGTAGTGTGTAAGTT	1740
DB	1756	GCGGATGCTCTCTTTGTTGTCCAAAAGAACCTGGAGAACATCCCGTAGTGTGTAAGTT	1815
QY	1741	CCGGTGGCTACTGTGGGCCACGTGTCCTGCCATGTTTCCCAAGTCTGCCTGGGCATCC	1800
DB	1816	CCGGTGGCTACTGTGGGCCACGTGTCCTGCCATGTTTCCCAAGTCTGCCTGGGCATCC	1875
QY	1801	TGTGTGTGGCACCAGCCCACTGGAGTGACTGCGCTCACTCCCGGAGGATGACATAGAG	1860
DB	1876	TGTGTGTGGCACCAGCCCACTGGAGTGACTGCGCTCACTCCCGGAGGATGACATAGAG	1935
QY	1861	GTGGTGGTGACCCAGGGGCCATTGGGGCTTATGCTTGGAGTGTCAGGAGGTGGAGCC	1920
DB	1936	GTGGTGGTGACCCAGGGGCCATTGGGGCTTATGCTTGGAGTGTCAGGAGGTGGAGCC	1995
QY	1921	GCCCGGTGGTGCTCTTATAGCTTGGTGGGGCAGCCAGCGGGACCCCTCAAAACCGG	1980
DB	1996	GCCCGGTGGTGCTCTTATAGCTTGGTGGGGCAGCCAGCGGGACCCCTCAAAACCGG	2055
QY	1981	GCCACACACCTGTGGGGGCTGGATTGGTTGGCTTTCTCTGGGTGTTCTTGCGAGCATCC	2040
DB	2056	GCCACACACCTGTGGGGGCTGGATTGGTTGGCTTTCTCTGGGTGTTCTTGCGAGCATCC	2115
QY	2041	CTCACTCTCTCTGTATTGTCGCGCTCAGCAGCGTCGGGCAGAGGGAGCTTCTAGCT	2100
DB	2116	CTCACTCTCTCTGTATTGTCGCGCTCAGCAGCGTCGGGCAGAGGGAGCTTCTAGCT	2175
QY	2101	AGACAAAGGTGGGCTTAGATCTGGGGGCTCCACCTTCTGGGACCAACAGCTATAGTCAG	2160
DB	2176	AGACAAAGGTGGGCTTAGATCTGGGGGCTCCACCTTCTGGGACCAACAGCTATAGTCAG	2235
QY	2161	GACCTCTCTCTCTGCGCTGAAGATGAACGGCTGCCCTGGCCCTGGGTAGCGGGGC	2220
DB	2236	GACCTCTCTCTCTGCGCTGAAGATGAACGGCTGCCCTGGCCCTGGGTAGCGGGGC	2295
QY	2221	AGTGGTTTTGGTGGCTTCCCTCCACCTTCTCTGCTGGATTCTTGGCCCAAGCCAGGCCAC	2280
DB	2296	AGTGGTTTTGGTGGCTTCCCTCCACCTTCTCTGCTGGATTCTTGGCCCAAGCCAGGCCAC	2355
QY	2281	ATCCGGCTCACTGGGGCGCCTCTAGCCACGTGTGATGAGACCTCCATCTAA	2331
DB	2356	ATCCGGCTCACTGGGGCGCCTCTAGCCACGTGTGATGAGACCTCCATCTAA	2406

RESULT 5

US-10-144-771-19015

US-10-144-771-19013
; Sequence 19015, Application US/10144771

GENERAL INFORMATION:									
; APPLICANT: VENTER, J. Craig									
; TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF									
; FILE REFERENCE: CL001321									
; CURRENT APPLICATION NUMBER: US/10/144.771									
; CURRENT FILING DATE: 2002-05-15									
; NUMBER OF SEQ. ID NOS: 47235									
; SEQ ID NO 19015									
; LENGTH: 2334									
; TYPE: DNA									
; ORGANISM: HUMAN									
US-10-144-771-19015									
Query Match 91.8%; Score 2138.8; DB 46; Length 2334;									
Best Local Similarity 95.1%; Pred. No. 0;									
Matches 2219; Conservative 0; Mismatches 112; Indels 3; Gaps 1;									
Qy	1	ATGCTTCCAGGGCCGAGCGCCCGCCGCGCCCGCGCGCTCTCGGTCTTTCCCTTC	60						
Db	1	ATGCTGGCAGGGCCGAGCGCCCGCCGCGCGCTCTCGGTCTCTCTCTTC	60						
Qy	61	CCGCCGCGGTGTC---GCTGCTGCTGCTGCTGGCGATACTAAGCGCCCGGTGTGCGC	117						
Db	61	CCGCCGCGGTGCTGCTGCTGCTGCTGCTGGCGATGCTGAGCGCCCGGTGTGCGC	120						
Qy	118	CGGTCCCGCTCAGTGCCTCCAGAACCTCGCTGCCCATCTCCGAGGCTGACTCTATCTC	177						
Db	121	CGGTCCCGCTCGGTGCCAGAACCTCGCTGCCCATCTCCGAGGCTGACTCTATCTC	180						
Qy	178	ACCGGTTTGCAGCGTCTCATACGTACAATTACTGCTGCTCTCTGTTGGATCTCGCTCC	237						
Db	181	ACCGGTTTGGCGCCCTCATACGTACAATTACTGCTCTCTCTGTTGGATCTCGCTCT	240						
Qy	238	CACACACTTTACGTGGTGACGGGATAGCATTTTCGCTTTAAACCTCCCGCTCTCTGGG	297						
Db	241	CACACACTTTACGTGGTGACGGGACAGCATTTTCGCTTTAAACCTCCCGCTCTCTGGG	300						
Qy	298	GAAGACCCCGAAGGATCGACTGATGGTACCTGAGACTCACACAGAACTGCAGGAAG	357						
Db	301	GAAGAGCCTCGAAGGATCGACTGATGGTGCCGAGACTCACACAGAACTGTAGGAAG	360						
Qy	358	AAAGCAAGAAAGAGGACGAATGCACAATTTATCCAGATTCTCGCCATTGTCAATGCC	417						
Db	361	AAAGCAAGAAAGAGGACGAATGCACAATTTATCCAGATTCTCGCCATTGCCAATGCC	420						
Qy	418	TCTCACCTCTCACGTGCGGCACCTTCGCTTTTGATCCGAAGTGGGGGTTATTGATGTG	477						
Db	421	TCTCACCTCTCACGTGCGGCACCTTCGCTTTTGATCCGAAGTGGGGGTTATTGATGTG	480						
Qy	478	TCCAGTTTCCAGCAGGTGAAAGACTTCAGAGCGCGCGGGAATATGCTTTTGGAGCCA	537						
Db	481	TCCAGTTTCCAGCAGGTGAAAGACTTCAGAGTGGCGGGGGAATATGCTTTTGGAGCCA	540						
Qy	538	GCTCAACGGTCAGCAGCTGTAAATGGCTGGGGGCTCTCTACACCGCCACTGTGAAGAAC	597						
Db	541	GCTCAACGGTCAGCAGCTGTAAATGGCTGGGGGCTCTCTACACCGCCACTGTGAAGAAC	600						
Qy	598	TTCTTGGGACTGAGCCCATCATCTCCGAGCTGTGGTGCAGCTGAGGACTGGATTCGA	657						
Db	601	TTCTTGGGACAGACCGGATTTATCTCCGAGCTGTGGTGCAGCTGAGGACTGGATTCGA	660						
Qy	658	ACAGAGACCTTGTATCTGGCTTAATGCTCCACGCTTGTGCGAGCTATGCTCCTGAGC	717						
Db	661	ACAGAGACCTTGTATCTGGCTTAATGCTCCACGCTTGTGCGAGCTATGCTCCTGAGC	720						
Qy	718	CCAGCTGAGTGGGGGATGAAGATGGAGACGATGAAATCTTTTCTTCCACGAGACC	777						
Db	721	CCGCTGAGTGGGGGATGAAGATGGAGACGATGAAATCTTCTTCCACGAGACC	780						
Qy	778	TCCGAGTGTGGACTCCTATGAGCGCATCAAGGTCCCAAGAGTGGCCCGGAGTGTGTGG	837						
Db	781	TCCGAGTGTGGACTCCTATGAGCGCATCAAGGTCCCAAGTGGCCCGGAGTGTGTGG	840						

Qy	721	GCTGAGTGGGGGATGAAGATGAGACGATGAATCTTTTTTTCTTTCACGGAGACCTCC	780
Db	808	GCCGAATGGGGGATGAAGATGAGACGACGAATCTACTTCTTTTACGGAGACTTT	867
Qy	781	CGAGTGTGGACTCCTATGAGCGCATCAAGGTGCCAAGAGTGGGCCCGAGTGTGTGCGGGG	840
Db	868	CGAGCATTTGACTCATACGAGCGCATTTAAAGTCCACGGGTGGCCCGTGTGTGTCGCGGG	927
Qy	841	GACCTTGGGGCAGGAAGACCCTTACGACAGAGATGACGACGTTTCTTGAAGGCTGACCTG	900
Db	928	GACCTCGGGGCGGAAGACCCCTCCACGACAGATGGACGACGTTTTTTGAAAGCTGACCTG	987
Qy	901	CTGTGCCACAGGGCCGAGCATGCGCGGGCCTCGGGGTTCTGCAAGGCTATGGCAGACGTT	960
Db	988	CTCTGTCCAGGGCTGAGCATGGCCGGGCCCTCCAGTGTCTGCAAGGATGTTGTGTGCTTT	1047
Qy	961	CGGCCCTCAGCCTGGAGCGGAACCCCATCTTTTATGGGATCTTTTCTCCCACTGGGAA	1020
Db	1048	CGACCTCAGCTTGGGGCAGGAGCTCCCATCTTTTATGGCATCTTTTCTCCCACTGGGAG	1107
Qy	1021	GGAGCTGCCATCTCTGCTGTGTGTCCTCCGACCCCAAGACATCCGGGAGTGCTGAAT	1080
Db	1108	GGGGCTACTATCTCTGCTGTGTGTCCTTCGGACCACAAGACATTCGACAGACTGCTGAAT	1167
Qy	1081	GGTCCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCGTGTCACTGGACAAAGAG	1140
Db	1168	GGTCCCTTCAGAGAACTAAACATGACTGCAACAGGAGACTGCGCTGCTGGACAATGAT	1227
Qy	1141	GTGCCCCAGCCACAGACCTGGAGAGTGCATCGCCCAACATGAAGCTCCACGACGTTTGGGA	1200
Db	1228	GTGCCCCAGCCACAGACTGGAGAGTGCATCACCACAACATGAAGCTCCGGCAGCTTTGGC	1287
Qy	1201	TCCTCACTCTCCCTGCCAGACCGGCTGCTCACCCTTATCAGAGACCAACCCTCTCATGGAC	1260
Db	1288	TCATCTCTCTCCCTGACCCGGTACTCACCTTTCATCGGGAGCACCACTCATCGAC	1347
Qy	1261	AGGCCGCTGTTCCCGGCTGACGGCCGCCCTCTGCTGTCTACTACAGATACAGCCTATCTC	1320
Db	1348	AGGCCAGTGTTTCCAGCTGATGGCCACCCCTCTGCTGTCTACTACAGATACAGCCTATCTC	1407
Qy	1321	AGAGTCTGGCCACACAGGCTGACCGACCTCTCAGGGAAAGNATATGACGTGCTCTACCTG	1380
Db	1408	AGAGTCTGGCCACACAGGCTGACCGACCTCTCAGGGAAAGAGTATGATGTCTCTACCTG	1467
Qy	1381	GGGACAGAGATGGACACCTCCACCGGGCTGTGGGCATGGAGCTCAGCTCAGTGTCTTG	1440
Db	1468	GGGNACAGAGATGGACACCTCCACCGAGCAGTGGGATCGGAGCTCAGCTCAGCGTCTT	1527
Qy	1441	GAGGATCTGGCCTTGTTCAGAACACAGCCGCTTGAGAGCATGAAATGTATACACGAT	1500
Db	1528	GAAGATCTGGCCTTATTTCCAGAGCCACAGCCAGTTGAGAACATGAAATGTATACACAGC	1587
Qy	1501	TGGCTCTGGTGGGCTCCCATACTGAGGTGACACAAAGTGAACACCACTGTGGCCGT	1560
Db	1588	TGGCTCTGGTGGGCTCCCATCTGAGGTGACACAAAGTGAATACAAACCACTGTGGCCGT	1647
Qy	1561	CTCCAGAGCTGCTCGGAGTGTATCCTGGCCACAGACCCGCTGTGGCCTGGAGCTCCCG	1620
Db	1648	CTCCAGAGCTGCTCAGAGTGCATCTCGCCACAGACCCGCTGTGGCTGGAGCTCCCG	1707
Qy	1621	CTTGATCTCTGTGTGGCCACCGCGGAGCACCGCGGGATGGTTCAAGATATAGAGTCA	1680
Db	1708	CTGGATCAGTGTGTGGCCCATGCCGGGAGCACCCAGGGTTGGTCCCAAGACATAGAGTCA	1767
Qy	1681	CGGATCTCTCTCTTGTGTCCAAAGAACCTTGGAGAACATCCCGTAGTGTGTTGAAGTT	1740
Db	1768	GCAGATCTCTCTCTTGTGTCTTAAAGACCTTGGAGAAGCTCCAGTAGTGTGTTGAAGTT	1827
Qy	1741	CCGGTGGCTACTGTGGGCCAGTGGTCTTGCATGTTCCCGAGTTCTGCTCGCTGGGCATCC	1800
Db	1828	CCGGTGGCTACAGCTGCGCATGTGGTCTTGGCATGTTCTCCAAGTCTCAGCATGGGCATCC	1887
Qy	1801	TGTTGTGTGGACACCGCCAGTTGGAGTGACTCGCTCACTCCCCGGAGGGATGGACTAGAG	1860

Db	1108	GGGCTACTATCTCTGCTGTCTGTGCCCTTCGGACCACAGACATTCGGACAGTGTGAAT	1167
Qy	1081	GGTCCCTTTAGAGAGCTAAACATGACTGCAACAGGGGAGTGCCTGTCTATGGACAACGAG	1140
Db	1168	GGTCCCTTCAGAAACTAAACATGACTGCAACAGAGGACTGCTGTGCTGGACAATGAT	1227
Qy	1141	GTGCCCAGCCCGACACCTGGAGAGTGCATCGCCAAACAATGAAGCTCCAGCAGTTTGG	1200
Db	1228	GTGCCCAGCCCGACACCTGGAGAGTGCATCACCACAACATGAAGTTCGGCGACTTTGGC	1287
Qy	1201	TCCTCACTCTCCCTGCCAGACCGGCTGCTCACCTTTATCAGAGCACACCCTCTCATGGAC	1260
Db	1288	TCATCTCTCCCTGCCTGACCGGGTACTCACCTTATCCGGGACCACCACCTCATGGAC	1347
Qy	1261	AGGCCCCGTTCGGGTGACGGCGCCCTGCTGCTCACTACAGATACAGCCTATCTC	1320
Db	1348	AGGCCAGTGTTCACAGCTGATGGCCACCCCTGCTGTCTACTACAGATACAGCCTATCTC	1407
Qy	1321	AGAGTCTGTGCCACAGGGTGACCAGCCTCTCAGGGAAAGAAATGACGTGCTCTACCTG	1380
Db	1408	AGAGTCTGTGCCACAGGGTGACCAGCCTCTCAGGGAAAGAGTATGATGTGCTCTACCTG	1467
Qy	1381	GGGACAGAGATGGACACCTCCACCGGGCTGTGCGCATTTGAGCCTCAGCTCAGTGTCTTG	1440
Db	1468	GGGACAGAGATGGACACCTCCACCGAGCAGTGGGATCGAGCTCAGCTCAGCTGCTT	1527
Qy	1441	GAGGATCTGGCCTGTTCCCAGAACACACACCGCGTTGAGAGCATGAATTTGTACCAAGAT	1500
Db	1528	GAGATCTGGCCTTATTCACAGAGCCACACGCCAGTTGAGAACATGAATTTGTACCAACAGC	1587
Qy	1501	TGGCTCCTGTGGGCTCCCATCTACTGAGGTGACACAAGTGAACACCACTCTGGCCGT	1560
Db	1588	TGGCTCCTGTGCTCCCGTACTGAGTGCACACAGTGAATACACCACTCTGGCCGT	1647
Qy	1561	CTCCAGAGCTGCTCGGAGTGTATCTGTGGCCAGACCCCTGTGTGCGCCTGGAGCTTCCGG	1620
Db	1648	CTCCAGAGCTGCTCAGAGTGCATCTGTGCCACAGACCCAGTCTGTGCTGGAGCTTCCGG	1707
Qy	1621	CTTGATCTTGTGTGGCCACGCGCGGAGCACCGGGGATGTTCAAGATATAGAGTCA	1680
Db	1708	CTGGATGAGTGTGTGGCCATGCGGGGAGCACCGAGGGTTGGTCCNAGACATAGAGTCA	1767
Qy	1681	CGGGATGCTCTCTTTGTGTCCAAAAGAACCTGGAGAACATCCCTGAGTGTGTTGAAGTT	1740
Db	1768	GCAGATGCTCTCTTTGTGTCTTAAAGAGCCTGGAGAGCTCCAGTGTGTTGAAGTT	1827
Qy	1741	CCGTGCTCTACTGTGGCCACAGTGTCTGCCATGTCCCCAGTGTCTGCTCGGCATCC	1800
Db	1828	CCGTGCTCTACAGCTGGCATGTGGTCTTGCCATGTTCTCCAAGTCAGCATGGGATCC	1887
Qy	1801	TGTGTGTGGCACCAGCCAGTGGAGTGACTGCGCTCACTCCCGGAGGGATGACATAGAG	1860
Db	1888	TGTGTGTGGCACCAGCCAGTGGAGTGACTGCTACACCCCGCGGGATGACATGGAG	1947
Qy	1861	GTGGTGTGACCCAGGGGCCATGGGGCTTATGCTTGGAGTGTGAGAGGTGGAGCC	1920
Db	1948	GTGGTGTGACCCAGGGGCCATGGGGCTTATGCTTGGAGTGTGAGAGGTGGGGCA	2007
Qy	1921	GGCCGCTGTGTGCTTATAGCTTGTGTGTGGGAGCCAGCGGGACCCCTCAAAACCGG	1980
Db	2008	GCCATGTGTAGCAGCTTACAGTGGTATGGGAGCCAGCGAGATGCTCGAGCCGG	2067
Qy	1981	GCCACACACCTTGTGGGGCTGGATGTGGTGTCTTCTGTGGGTGTCTGTGAGCATCC	2040
Db	2068	GCCACACACACCTTGTGGGGCTGGATGTGGTGTCTTCTGTGGGTGTCTGTGAGCATCC	2124
Qy	2041	CTCACTCTCTCTGTGTGTGCTGCGCTGACGACGCTGCGGACAGAGGAGCTTCTAGCT	2100
Db	2125	CTGACTCTCATTTCTGATTTGTGCGGCTCAGCAGCGGCGCCACAGAGGAACTTCTGGCT	2184
Qy	2101	AGAGACAAGTGTGGCTTAGATCTGGGGCTTCCACCTTCTGGGACCAACAAGCTATAGTCA	2160

Db	2185	AGACAAAGTGGCGCTGGACCTGGGGGCTCCACCTTCTGGGACCACAAAGCTACAGCCAA	2244
Qy	2161	GACCTCCCTCTCCTTCGCCTGAAGATGAACGGCTGCCCTGGCCCTGGGTAAAGGGGC	2220
Db	2245	GACCTCCCTCCCTCTCCTTGAAGATGAGCGTTCGCCCTGGCCCAAGAGGGGC	2304
Qy	2221	AGTGGTTTGGTGGCTTCCTCCACCCTTCTGATTCCTTGGCCCAAGCCAGCCAC	2280
Db	2305	AGTGGCTTTGGTGATTCTCACCACCTCTCCTGCTGATCTTGGCCCAAGCCAGCCAC	2364
Qy	2281	ATCCGGCTCACTGGGGCGCTCTAGCCACCTGTGATGAGACCTCCACTA	2330
Db	2365	ATTCGGCTAACTGGGGCTCTTAGCCACATGTGATGAACATCCACTA	2414
RESULT 9			
US-10-170-235-32596			
; Sequence 32596, Application US/10170235			
; GENERAL INFORMATION:			
; APPLICANT: VENTER, J. Craig			
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY			
; FILE REFERENCE: CL001380			
; CURRENT APPLICATION NUMBER: US/10/170,235			
; CURRENT FILING DATE: 2003-03-17			
; NUMBER OF SEQ ID NOS: 42514			
; SEQ ID NO 32596			
; LENGTH: 3641			
; TYPE: DNA			
; ORGANISM: HUMAN			
US-10-170-235-32596			
Query Match 78.4%; Score 1827.6; DB 47; Length 3641;			
Best Local Similarity 87.3%; Pred. No. 0;			
Matches 203; Conservative 0; Mismatches 279; Indels 18; Gaps 2;			
Qy	1	ATGCTTCCAGGCGGAGCGGCCCGCGCGGGCCCGCGCGGCTCCGGTCTTTCCCTTC	60
Db	87	ATGCGGGCTCTGCTGGCGGCGCGCGCGGGTCCGGGCAGCCTACAGCTCGCCCTTC	146
Qy	61	CGCGCGCGCTGCTCGCTGCTGCTGGCGATACTAAGCGCCCCGGTGTGCGGCGC	120
Db	147	CC-----GCTACTGCTGCTGGCGGTGCTGAGCGGCGCGGTATCCGGCCGC	191
Qy	121	GTCCCGCGCTCAGTGCCGAGAACCTCGCTGCCCATCTCGAGGCTGACTCTATCTCAC	180
Db	192	GTCCCGCGCTCGGTGCCGAGAACCTCGCTTCCAATCTCTGAGGCTGACTCTCTCAC	251
Qy	181	CGGTTTCACGGTCTCATAGCTACAATTACTCTGCTCTCCTGTGTGATCCTGCTCCC	240
Db	252	CGGTTTCGACGTCTCTACACATACAATTACTCTGTTCTCCTGTGTGATCCTGCTCCC	311
Qy	241	ACACTTACCTGCGTGCACGGGATAGCATCTTCGCTTTAAACCTTCTCTGCGGAA	300
Db	312	ACACTTATGTTGGCGCGGGACACCATCTTCGCTTTATCCCTGCGCTTCTCAGGGAG	371
Qy	301	AGACCCGAGGATCGACTGGATGGTACCTGAGACTCACAGACAACTGCAGGAAGAA	360
Db	372	AGACCCGCGAGGATTGACTGGATGGTTCTCTGAGGCTCACAGACAGACTGTAGGAAGAA	431
Qy	361	GGCAAGAAAGACGAATCTACAATTTATCCAGATCTCGCCATCTGCAATGCTCT	420
Db	432	GGCAAGAAAGACGAATGTACAATTTGTCCAGATCTTCGCCATTCGCAATGCTCT	491
Qy	421	CACCTCTACGTCGCGACCTTCGCTTTTGATCCGAAGTCGGGGTTATTGATGTGCC	480
Db	492	CACCTCTACTTGTGCGACCTTCGCTTTTGATCCGAAGTCGGGGTTATTGATGTGCC	551
Qy	481	AGTTTCAGCAGGTTGAAAGACTTGAGCGGCGGGGGAATGTCTTTGAGCAGCT	540
Db	552	AGTTTCAGCAGGTTGAAAGACTTGAGATGCGCGGGGGAATGTCTTTGAGCAGCT	611
Qy	541	CAACGGTCAACAGCTGTAAATGGCTGGGGCGTCTCTACACGCCACTGTGAAGAACTTC	600

APPLICATION NUMBER: EP 98300693.3
FILING DATE: 30-JAN-1998
APPLICATION NUMBER: US 9816423.9
FILING DATE: 28-JUL-1998
ATTORNEY/AGENT INFORMATION:
NAME: Prestia, Paul F.
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GP-30036
TELEPHONE: 601-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2893 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-160-762-1

Query Match 69.8%; Score 1626; DB 16; Length 2893;
Best Local Similarity 83.2%; Pred. No. 0;
Matches 1938; Conservative 0; Mismatches 275; Indels 117; Gaps 3;

QY	1	ATGCTTGCAGGGCCGAGCGCCCGCGGGCCCGCGCGCTCCGGTCTCTTCCCTTC	60
Db	110	ATGCGCGGCTCTGCTGCGGGCCCGCGGGTCCCGGCGAGCTACAGCCCTCGCCCTTC	169
QY	61	CGCGCGCGCTGTGCGTGTGCTGCTGCGGCGGATTAAGCGCCCGCGTGGCGCGC	120
Db	170	CC-----GCTACTGCTGCTGGCGGCTGCTAGCGCGCGCGGATCCGGCGCGC	214
QY	121	GTCCCGCGCTAGTGCAGAGACCTCGCTGCCATCTCCGAGGCTGACTCTCTACCC	180
Db	215	GTCCCGCGCTCGTGCAGAGACCTCGCTTCCAATCTCTGAGCTGACTCTCTCACCC	274
QY	181	CGGTTTGCAGGCTCATAGTACAAATTAATCTGCTCTCTCTGTGTGATCCTGCTCCAC	240
Db	275	CGGTTTGCAGGCTCATAGTACAAATTAATCTGCTCTCTCTGTGTGATCCTGCTCCAC	334
QY	241	ACACTTTAGTCTGGTGCAGCGGATAGCATCTTCGCTTTAAACCTCCCTCTCTGGGAA	300
Db	335	ACACTTTAGTCTGGTGCAGCGGATAGCATCTTCGCTTTAAACCTCCCTCTCTGGGAA	394
QY	301	AGACCCGAGGATCGACTGGATGATGCTGAGACTCACAGACAGAACTGCAGGAGAAA	360
Db	395	AGACCCGAGGATCGACTGGATGATGCTGAGGCTCACAGACAGAACTGTAGGAGAAA	454
QY	361	GGCAAGAAAGAGGACCAATGTCACAAATTTATCCAGATTCTCGCCATTTGTCATGCTCT	420
Db	455	GGCAAGAAAGAG-----	466
QY	421	CACCTCTCACGTGGCGACCTTCGCTTTTGATCCGAGTGGCGGTTATTGATGTGCC	480
Db	467	-----GATGTGTC	475
QY	481	AGTTTCCAGCAGGTTGAAAGACTTGAGACGCGCGGGGAAATGTCTTTTTCAGCCAGCT	540
Db	476	AGTTTCCAGCAGGTTGAAAGACTTGAGAGTGCGCGGGGAAATGTCTTTTTCAGCCAGCT	535
QY	541	CAACGGTTCAGAGCTGTAAATGGCTGGGGCGTCTCTACACCGCACTGTGAAGACTTC	600
Db	536	CAGCGTTCAGAGCTGTAAATGGCTGGGGCGTCTCTATGCTGCCACTGTGAAGAACTAC	595
QY	601	CTGGGAGTACGCCATCATCTCCGAGCTGGGTGCGAGCTGAGGACTGGATTTCGAACA	660
Db	596	CTGGGAGTACGCCATCATCTCCGAGCTGGGTGCGGAGGACTGGATTTCGAGACA	655
QY	661	GAGACCTTGTCTCTGCTTAATGCTCCAGGCTTTGTGCGAGCTATGGTCTCTGAGCCCA	720
Db	656	GATACCTTGCCTTCTGCTGAACGCCCGAGCTTTGTGCGAGCGCTGGCTTTGAGCCCA	715

QY	721	GCTGAGTGGGGGATGAAGATGAGACGATGAATCTTTTCTTCTACGGAGACCTCC	780
Db	716	GCCGAATGGGGGATGAAGATGAGACGATGAATCTTCTTCTTACGGAGACTTC	775
QY	781	CGAGTGTTCGACTCTATGAGCGCATCAAGGTCCCAAGAGTGGCCCGAGTGTGTGGGG	840
Db	776	CGAGCATTTGACTCATACGAGCGCATTAAGTCCCGGCTGGCCCGTGTGTGTGGGG	835
QY	841	GACCTTGGGGGAGGAGACCTTTCAGCAGAGATGACGACGCTTTCTGAAGGCTGACCTG	900
Db	836	GACCTTGGGGGAGGAGACCTTTCAGCAGAGATGACGACGCTTTTGAAGCTGACCTG	895
QY	901	CTGTGCCAGGCGCGAGCATGGCGGGCTCCGGGTTCTGAGGCTATGGCAGAGCTT	960
Db	896	CTGTGCCAGGCGCTGAGCATGGCGGGCTCCAGTGTCTGAGGATGTGTGTGCTT	955
QY	961	CGGCTCAGCTGGAGCGGGAACCCCATCTTTTATGGGATCTTTTCTCCAGTGGAA	1020
Db	956	CGACCTGAGCTTGGGCGAGGACTCCCATCTTTTATGGGATCTTTTCTCCAGTGGAG	1015
QY	1021	GGAGTGTCCATCTCTGCTGTGTGCTTCCGACCCCAAGAGATCCGGGAGTGTCTGAAT	1080
Db	1016	GGGCTACTATCTGCTGTGCTTCCGACCAACAGACATTCGGACAGTGTCTGAAT	1075
QY	1081	GGTCCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCTGTCTATGACAGAG	1140
Db	1076	GGTCCCTTTAGAGAGCTAAACATGACTGCAACAGAGAGTGCCTGTCTGAGCAATAT	1135
QY	1141	GTGCCCCAGCAGCTGAGAGTGCATCGCCCAACAACTGAAGCTCCAGGATTTGGA	1200
Db	1136	GTGCCCCAGCAGCTGAGAGTGCATCGCCCAACAACTGAAGCTCCAGGATTTGGA	1195
QY	1201	TGCTCACTCTCCCTGCGAGACCGGCTGCTACCTTTTATCAGAGACCACTCTCTATGGAC	1260
Db	1196	TGCTCACTCTCCCTGCGAGACCGGCTGCTACCTTTTATCAGAGACCACTCTCTATGGAC	1255
QY	1261	AGGCGGCTTCCCGGCTGAGCGCGCCCTGCTGTGGTCACTACAGATACAGCTATCTC	1320
Db	1256	AGGCGGCTTCCCGGCTGAGCGCGCCCTGCTGTGGTCACTACAGATACAGCTATCTC	1315
QY	1321	AGAGTGTGGCGCCACAGGCTGACCGCTCTCAGGAAAGAAATATACGCTGTCTACCTG	1380
Db	1316	AGAGTGTGGCGCCACAGGCTGACCGCTCTCAGGAAAGAAATATACGCTGTCTACCTG	1375
QY	1381	GGGACAGAGGATGGACACCTCCACCGGCTGTGCGCATTTGGAGCTCAGCTAGCTTTC	1440
Db	1376	GGGACAGAGGATGGACACCTCCACCGGCTGTGCGCATTTGGAGCTCAGCTAGCTTTC	1435
QY	1441	GAGGATCTGGCCTTGTTCACAGAACACAGCGGTTGAGAGCATGAATTTGACACAGAT	1500
Db	1436	GAGGATCTGGCCTTGTTCACAGAACACAGCGGTTGAGAGCATGAATTTGACACAGAT	1495
QY	1501	TGGCTCTCTGGTGGGCTCCCATACTGAGGTGACACAGTGAACACCAAGCACTGTGGCGT	1560
Db	1496	TGGCTCTCTGGTGGGCTCCCATACTGAGGTGACACAGTGAACACCAAGCACTGTGGCGT	1555
QY	1561	CTCCAGAGCTGCTGGAGTGTATCTTGGCCCGGAGGACCCCGTGTGGCCCTGGAGCTTCCGG	1620
Db	1556	CTCCAGAGCTGCTGGAGTGTATCTTGGCCCGGAGGACCCCGTGTGGCCCTGGAGCTTCCGG	1615
QY	1621	CTTGTATGTTGTGGCCCGGAGCAGCAGCGGGGATGGTTCAAGATATAGAGTCA	1680
Db	1616	CTTGTATGTTGTGGCCCGGAGCAGCAGCGGGGATGGTTCAAGATATAGAGTCA	1675
QY	1681	GGGATGTCTCTTGTGTCCAAAGAACCTGGAGAACATCCCGTACTGTGTGAAGTT	1740
Db	1676	GGGATGTCTCTTGTGTCCAAAGAACCTGGAGAACATCCCGTACTGTGTGAAGTT	1735
QY	1741	CGGTTGGCTTACTGTGGGCGGAGCTGTGCTTCCGCTTCTCCCGGAGTCTCTCCCTGGGATCC	1800
Db	1736	CGGTTGGCTTACTGTGGGCGGAGCTGTGCTTCCGCTTCTCCCGGAGTCTCTCCCTGGGATCC	1795
QY	1801	TGTGTGTGGCAGCCCGAGTGTGCTGCTCACTCTCCCGGAGGATGGACTAGAG	1860

```
Db 1796 TGTGTGGACACAGCCAGTGGAGTGAATCACTCACTACCCCGCGGGATGGACTGGAG 1855
Qy 1861 GTGGTGTGACCCAGGGGCCATGGGGCTTATGCTTGCAGTGTCTCAGGAGGTGGAGCC 1920
Db 1856 GTGGTGTGACCCAGGGGCCATGGGGCTTATGCTTGTGAATGTCAAGAGGGTGGGGCA 1915
Qy 1921 GCGCGGTGTGGTGTCTATAGCTTGTGTGGGCGAGCCAGCGGGACCCCTCAACACGG 1980
Db 1916 GCCCATGTGGTAGCAGCTTACAGCTTGGTATGGGCGAGCCAGAGAGTGTCCGAGCGG 1975
Qy 1981 GCCCACACCGTGTGGGGGCTGGATGGTGGCTTTCTCCTGGGTGTCTTGCAGCATCC 2040
Db 1976 GCCCACAC- --AGTGGGGGCGGGACTGGCTGGCTTCTTGGGGATTTCTCGCAGCATCC 2032
Qy 2041 CTCACTCCCTCCCTCTCTCTGATGTGGCTGCGGTGAGAGCGGTGGGCGACAGAGGGAGCTTAGCT 2100
Db 2033 CTGACTCTCATCTCTGATGTGGCTGCGGTGAGAGCGAGCGGCGACAGAGGAACTTCTGGCT 2092
Qy 2101 AGAGACAAGTGGCTTAGATCTGGGGCTCCACCTTCTGGGACCAAGCTATAGTCAG 2160
Db 2093 AGAGACAAGTGGGCTTGGACCTGGGGCTCCACCTTCTGGGACCAAGCTACAGCAA 2152
Qy 2161 GACCTCGCTCTCTTGGCTTGAAGTGAAGCGGTGCCCTGGCCCTGGGTAAAGGGGGC 2220
Db 2153 GACCTCGCTCTCTCTCTGATGAAGTGAAGCGGTGGCGCTGGCGCTGGCCCAAGAGGGC 2212
Qy 2221 AGTGGTGTGGTGTCTCCCTCCACCTTCTGCTGGATCTTGGCCAGCGCCAGCCAC 2280
Db 2213 AGTGGCTGTGGTGTCTCAACACCTTCTGCTGATCTTGGCCAGCGCCAGCCAC 2272
Qy 2281 ATCGGGTCACTGGGGGCGCTTAGCCACGCTGTGATGAGACCTCATCTA 2330
Db 2273 ATTCGGCTAATGGGGCTCTCTAGCCACATGTGATGAACATCATCTA 2322
```

RESULT 11
US-09-160-762-1

```
; Sequence 1, Application US/09160762A
; GENERAL INFORMATION:
; APPLICANT: David Michalovich
; APPLICANT: Trudy Rachel Doe
; APPLICANT: Philip David Hayes
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP-30036
; CURRENT APPLICATION NUMBER: US/09/160,762A
; CURRENT FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2893
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-160-762-1
```

```
Query Match 69.8%; Score 1626; DB 16; Length 2893;
Best Local Similarity 83.2%; Pred. No. 0;
Matches 1938; Conservative 0; Mismatches 275; Indels 117; Gaps 3;

Qy 1 ATGCTTGCCAGGGCCGAGCGCCCGCGCGCCCGCGCGCCCTCGCGGTCTTCCCTTC 60
Db 110 ATGCGGGCTCTGCTGGCGCGCCCGCGCGCGCGCGCGCTACAGCCTCGCGCTTC 169
Qy 61 CCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db 170 CC-----GCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 214
Qy 121 GTCCCGCGCTCAGTGCCAGAACCTCGCTGCGCCATCTCCGAGGTGACTCTATCTACC 180
Db 215 GTCCCGCGCTCGGTGCCAGAACCTCGCTTCCATCTCTGAGGTGACTCTGCTCACC 274
Qy 181 CGGTGTGAGCGCTCTATACCTACAAATACCTCTGCTGCTGCTGCTGCTGCTGCTG 240
```

```
Db 275 CGGTTCGAGTCCCTCACACATACAAATTAATCTCTGTTCTCTCTGTTGATCTCCCTCCAC 334
Qy 241 ACACCTTACGTGCGTGACGGATAGCATCTTGCCTTTAAACCTCCCTCTCTCTGCGGAA 300
Db 335 ACACCTTATGTTGGCGCCGGACACCATCTTGCCTTTATCCCTGCCCTTCTCAGGGAG 394
Qy 301 AGACCCGAGATCGACTGGATGGTACCTCAGACTCACAGACAGAACTCAGGAGAGAA 360
Db 395 AGACCCGAGATGACTGGATGGTCTCAGGCTCACAGACAGAACTGTAGGAAGAA 454
Qy 361 GGCAAGAAGAGGAGGAATGTACAAATTTTATCAGATTTCTCGCCATTTGTCAATGCCCTCT 420
Db 455 GGCAAGAAGAG----- 466
Qy 421 CACCTCTCACGTGCGGCACCTTCGCTTTTGATCCGAAGTGGGGGTTATTTGATGTGTCC 480
Db 467 -----GATGTGTCC 475
Qy 481 AGTTTCCAGAGTTGAAAGACTTGAGAGCGCGCGGGGAAATGTCTTTTGGACAGCT 540
Db 476 AGTTTCCAGAGTTGAAAGACTTGAGAGTGGCGGGGAAATGTCTTTTGGACAGCT 535
Qy 541 CAACGGTCAGCAGCTGTAATGGCTGGGGCGCTCTCTACACGGCCACTGTGAAGAACCTTC 600
Db 536 CAGCGGTGAGCAGCTGTAATGGCTGGGGGCTCTCTATGCTGCCACTGTGAAAACCTAC 595
Qy 601 CTGGGGACTGAGCCATCATCTCCGAGCTGTGGTGCAGCTGAGGACTGGATTCGAACA 660
Db 596 CTGGGGAGGAGCCAAATATCACAGACAGTGGTGTGCCGAGACTGGATTCGGA 655
Qy 661 GAGACCTTGTCTCTGCTTAACTCCAGCCTTTTGTGCGAGCTATGGTCTCTGAGCCCA 720
Db 656 GATACCTTGGCTCTCTGCTGACGCCCGCCAGCTTTGTCGAGCGCTGGCTTTGAGCCCA 715
Qy 721 GCTGAGTGGGGGATGAAGATGGAGACATGAATCTTTTCTTCTTCCAGGAGACTTC 780
Db 716 GCCGAATGGGGGATGAAGATGGAGACAGCAAAATCTACTTCTTCTTACGGAGACTTC 775
Qy 781 CGAGTGTGTGACTCTCTATGAGCGCATCAAGTGCACAGAGTGGCCGAGTGTGTCGGGG 840
Db 776 CGAGCATTTGACTCATAGAGCGCATTAAGTCCCGCGGTGGCCGTGTGTGTCGGGG 835
Qy 841 GACCTTGGGGCAGGAGAACCTTTCAGCAGAGATGGAGCAGCTTTTCTGAAGCTGACCTG 900
Db 836 GACCTGGGGCGCGAAGACCTCTCAGCAGAGATGGAGCAGCTTTTGTGAAAGCTGACCTG 895
Qy 901 CTGTGCCCAGGGCCGAGCATGGCGGCTCTCCGGGTCTCGAGGCTATGGCAGAGCTT 960
Db 896 CTGTGCCAGGGCTGAGCATGGCGGCTCCAGTGTCTGACAGATGTGTGCTGTGCTT 955
Qy 961 CGGCTCAGCTGGAGCGGGAACCCCATCTTTTATGGATCTTTTCTCCAGTGGGAA 1020
Db 956 CGACCTGAGCTGGGGCAGGAGCTCCCATCTTTTATGGCATCTTTCTCCAGTGGGAG 1015
Qy 1021 GGAGCTGCCATCTCTGCTGTGTGCTGCTTCCGACCCCAAGACATCCGGGCGAGTGTGAAT 1080
Db 1016 GGGCTACTATCTCTGCTGTCTGCTTCCGACCAAGACATTTCCGACAGTGTGTAAT 1075
Qy 1081 GGTCCCTTTAGAGAGCTAAAACATGACTGCAACAGGGGACTGCTGTCTATGGACAACAG 1140
Db 1076 GGTCCCTTCAGAGAACTAAAACATGACTGCAACAGAGAGACTGCTGTCTGGACAATGAT 1135
Qy 1141 GTGCCCCAGCCAGACCTGGAGAGTGCATGCCAACACATGAAGCTCCAGCAGTTTGA 1200
Db 1136 GTGCCCCAGCCAGACCTGGAGAGTGCATCAACAAACATGAAGCTCCGACACTTTGGC 1195
Qy 1201 TCCTCACTCTCCCTGCCAGACCGGTGCTCACTCTTATCAGAGACCACTCTCTCATGGAC 1260
Db 1196 TCATCTCTCTCCCTGAGCGGTACTCACTCTTATCCGGGACCACTCATGGAC 1255
Qy 1261 AGGCGCGTGTCCCGGTGAGCGGCGCCCTGCTGTGCTACTACAGATACAGCCTATCTC 1320
Db 1256 AGGCGAGTGTTCAGAGCTGATGGCGACCCCTGCTGCTCACTACAGATACAGCCTATCTC 1315
```

Qy	1321	AGAGTCGTGGCCACACAGGGTGACACAGCCTCTCAGGGAAAGAAATATGACGTGCCTACCTG	1380
Db	1316	AGAGTCGTGGCCACACAGGGTGACACAGCCTCTCAGGGAAAGAGTATGATGTGCTACCTG	1375
Qy	1381	GGGACAGAGGATGGACACACCTCCACCGGGTGTGCGCATTTGGAGCTCAGCTCAGTGTCTTG	1440
Db	1376	GGGACAGAGGATGGACACACCTCCACCAGCAGTGGGATCGAGCTCAGCTCAGCTACGGTTC	1435
Qy	1441	GAGGATCTGGCCTTTGCCAGAACACACGCCGGTTGAGAGCATGAATTTGTACCAGGAT	1500
Db	1436	GAAGATCTGGCCTTTATTTCCAGAGCACACGCCAGTTTTCAGAACATGAATTTGTACCAGC	1495
Qy	1501	TGGCTCTGTGGGGTCCCACTACTGAGGTGACACAAAGTGAACACCAGCAACTGTGGCGT	1560
Db	1496	TGGCTCTGTGGCTCCCGTACTGAGGTGACACAAGTGAATACAACCAACTGTGGCGT	1555
Qy	1561	CTCCAGAGCTGTGGGATGTATCTTGGCCACAGACCCCGTGTGCGCCTGGAGCTTCCCG	1620
Db	1556	CTCCAGAGCTGTCTCAGAGTGCATCTTGGCCACAGACCCAGTCTGTGCTGGAGCTTCCCG	1615
Qy	1621	CTTGTATCTTGTGTGCCACCGCGGACACCGCGGGATGCTCAAGATATAGAGTCA	1680
Db	1616	CTGGATGAGTGTGTGCCCATGCGGGAGCACCGAGGGTTGTTCCAAAGACATAGAGTCA	1675
Qy	1681	GCGGATGTCTCTTTTGTGTCCAAAGAACCTGGAGAACATCCCGTAGTGTGAAGTT	1740
Db	1676	GCAGATGTCTCTTTGTCTTAAAGACCTGGAGAACGTCACGTTAGTGTGAAGTT	1735
Qy	1741	CCGGTGGCTACTGTGGCCACGTGGTCTGCCATGTTCCCCAGTTCTGCCTGGGATCC	1800
Db	1736	CCCGTGGCTACAGCTGCGCATGTGGTCTTGCCCATGTTCTCCAAGCTCAGCATGGGATCC	1795
Qy	1801	TGTGTGTGGCACACAGCCAGCTGGAGTGACTTGCCTCAGTCCCGCGGAGGATGACATAGAG	1860
Db	1796	TGTGTGTGGCACACAGCCAGTGGAGTGACTTGCCTCAGTCAACCCCGCGGGATGACATGGAG	1855
Qy	1861	GTGGTGTGACCCAGGGGCCATGGGGCTTATGCTTTCGAGTGTTCAGGAGGTGGAGCC	1920
Db	1856	GTGGTGTGACCCAGGGGCCATGGGGCTTATGCTTTCGAGTGTTCAGGAGGTGGGGCA	1915
Qy	1921	GCCCGGTGTGTGCTGTATACCTTGGTGTGGGCGAGCCAGCGGGACCCCTCAAAACCGG	1980
Db	1916	GCCCATGTGTACAGCTTACACTTGGTATGGGCGACCGACGAGATGCTCCGAGCCGG	1975
Qy	1981	GCCCAACACCTTGTGGGGCTGGATTGGTTGGCTTCTCTCTGGGTGTTCGAGCATCC	2040
Db	1976	GCCCAACAC --- AGTGGGGCGGAGCTGGCTGGCTTCTTCTTGGGGATCTTCGAGCATCC	2032
Qy	2041	CTCACTCTCTCTGATTGTGCGCGTCAAGCTGAGCGTGGGACAGAGGGAGCTTCTAGCT	2100
Db	2033	CTGACTCTCTTCTGATTGTGCGCGTCAAGCTGAGCGTGGGCGAGAGGGAACTTCTGGGT	2092
Qy	2101	AGACAAGGTGGGCTTAGATCTGGGGGTCTCACTTCTGGGACCAACAGCTATAGTCAG	2160
Db	2093	AGACAAGGTGGGCTTGGACCTGGGGCTTCCACCTTCTGGGACCAACAGCTACAGCCAA	2152
Qy	2161	GACCTTCCCTCTCTTCGCTTGAAGATGAACGCTGCCCTGGCCCTGGGTGAAGCGGGC	2220
Db	2153	GACCTTCCCTTCCCTCTCTCTGAAGATGAGCGGTGGCCCTGGCCCTGGGCGCAAGAGGGC	2212
Qy	2221	AGTGGTTTTGGTGGCTTCCCTCCACCTTCTCTCTGGATTTGTCGCAAGCCAGCCAC	2280
Db	2213	AGTGGCTTGGTGGATTCTCACCACCTTCTCTCTGATCTTGGCCCAAGCCAGCCAC	2272
Qy	2281	ATCCGGCTACTGGGGCGCTCTAGCCAGTGTGATGAGACCTCCACTTA	2330
Db	2273	ATTCCGGCTACTGGGGCTCTCTTAGCCACATGTGATGAACATCCACTTA	2322

RESULT 12
US-09-808-665A-1

US 03 000 005A 1
; Sequence 1, Application US/09808665A

```
; GENERAL INFORMATION:
; APPLICANT: David Michalovich
; APPLICANT: Trudy Rachel Doe
; APPLICANT: Philip David Hayes
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP-30036-CI
; CURRENT APPLICATION NUMBER: US/09/808,665A
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 09/160,762
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 98300693.3
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 9816423.9
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: 9816676.2
; PRIOR FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2893
; TYPE: DNA
.- ORGANISM: HOMO SAPIENS
US-09-808-665A-1

Query Match      69.8%; Score 1626; DB 34; Length 2893;
Best Local Similarity 83.2%; Pred. No. 0;
Matches 1938; Conservative 0; Mismatches 275; Indels 117; Gaps 3;

Qy   1 ATGCTTGCCAGGGCGGCAGCGGCCCGCGCGGGCCCCGGCGGCTCGGTTCITTCGCTTTC 60
Db   110 ATGCGGGCCTCTGCTGCGGGCCCGCGCGGGTCCCGGGCAGCCTACAGCCTCGCCCTTC 169

Qy   61 CCGCGCGCGCTGCTGCGTGTGCTGTCTGGCGATACTAAGCCGCCCGGTGTGCGGCCGC 120
Db   170 CC-----GCTACTGCTGTGCGCGGTGTGAGCGGCGGTATCCGGCGCG 214

Qy   121 GTCCCCGGCTCAGTGCCCGAGAACCCTCGCTGCCCATCTCCGAGGCTGACTCTATCTCAC 180
Db   215 GTCCCCGGCTCGGTGCCAGAACCTCGTTCCAATCTCTGAGGCTGACTCTGTCTCAC 274

Qy   181 CGGTTTGCACGGTCTCATAGTACAATTACTCTGTCTCTCTTGTGGATCCTGCTCCCAC 240
Db   275 CGSTTCGCAGTCCCTCAACATACAATTACTCTGTCTCTCTTGTGGATCCTGCTCCCAC 334

Qy   241 ACATTTFACCTGCGTGCACGGGATAGCATCTTCGCTTTAACTCCCTTCTCTCTGGGGA 300
Db   335 ACATTTFATGTTGGCGCGGGACACATCTTCGCTTTATCCCTGCTCAGGGAG 394

Qy   301 AGACCCCGAAGGATCGACTGGATGGTAGCTGAGACTCACAGACAGAACTGCAGGAAGAA 360
Db   395 AGACCCCGAGGATTGACTGGATGGTTCTGTAGGCTCACAGACAGAACTGTAGGAAGAA 454

Qy   361 GGCAAGAAAGAGAGGAGNATGCAATTTTATCCAGATTCTCGCCATTTGTCAATGCTCT 420
Db   455 GGCAAGAAAGAG----- 466

Qy   421 CACCTCTCAGTGTGCGGCACCTTCGCTTTTGTATCCGAAGTGGCGGGTTATTGATGTGTC 480
Db   467 -----GAITGTTC 475

Qy   481 AGTTTCCAGCAGGTTGAAAGACTTGAAGCGCGCGGGGGAATAATGCTTTTGAGCCAGCT 540
Db   476 AGGTTCCAGCAGGTTGAAAGACTTGAAGTGGCGCGGGGGAATAATGCTTTTGAGCCAGCT 535

Qy   541 CAACGGTTCAGCAGCTGTAATGGCTGGGGGGCTCCTCTACACCGCCACTGTGAAGAAGTTC 600
Db   536 CAGCGGTTCAGCAGCTGTAATGGCTGGGGGGCTCCTCTATGCTGCCACTGTGAAAAACTAC 595

Qy   601 CTGGGACTCAGCCCATCATCTCCGAGCTGTGGGTGCGAGCTGAGACTGGATTCGAACA 660
Db   596 CTGGGACGAGGCCAATTATCACAGACAGTGGGTGCTGCCGAGGACTGGATTCGGA 655

Qy   661 GAGACCTTGTATCTCTGGCTTAATGTCTCCAGCCCTTTGTCCAGCTTTGTCCAGCTATG 720
```


Db 275 CGGTTCGAGTCCCTCACATACAAATTAATCTGTTCTCTCTGTTGATCCTGCCTCCAC 334
QY 241 ACACCTTACGTCGGTACAGGGATAGATCACTTTCGCTTTAAACCTCCCTCTCTCTGCGGAA 300
Db 335 ACACCTTATGTTGGCCCGGGACACCACTTCGCTTTATCCCTGCCTCTCAGGGGAG 394
QY 301 AGACCCCGAAGATCAGCTGGATGGTACCTGAGACTACAGACAGAACTGACAGGAAGAA 360
Db 395 AGACCCCGAGGATGACTGGATGGTTCCTGAGCTCACAGACAGAACTGTAGGAAGAA 454
QY 361 GGCACAAAGAGAGCAATGTACAAATTTATCCAGATTCTGCCAATTGTCAATGCCTCT 420
Db 455 GGCAGAGAG----- 466
QY 421 CACCTCTCACGTGCGGCACCTTCGCTTTTGTATCCGAAGTGGGGGTTATTGATGTGCC 480
Db 467 -----GATGTGCC 475
QY 481 AGTTTCCAGCAGTTGAAAGACTTGAGAGCGGCGGGGAAATGTCTTTTGTAGCCAGCT 540
Db 476 AGGTTCCAGCAGGTTGAAAGACTTGAGAGTGCGCGGGGAAATGTCTTTTGTAGCCAGCT 535
QY 541 CAACGGTCAGCAGCTGAATGCTGGGGGCTCCTCTACACGCCCACTGTGAAGACTTC 600
Db 536 CAGCGTCAGCAGCTGAATGCTGGGGGGTCTCTATGCTGCCACTGTGAAGAACTAC 595
QY 601 CTGGGACTGAGCCCATCATCTCCGAGCTGTGGGTGAGCTGAGACTGGATTCGAACA 660
Db 596 CTGGGACGAGCAATATACAGAGCAGTGGGTCTGTCGCGAGACTGGATTCGGACA 655
QY 661 GAGACCTGTCTATCTGGCTTAATGCTCAGCGCTTTGTGCGAGCTATGCTCTGAGCCCA 720
Db 656 GATACCTGCCCTTCTGGTGAACGCCCCAGCCTTTGCGCAGCGCTGGCCTTGAGCCCA 715
QY 721 GGTGAGTGGGGATGAAGATGAGACGATGAATCTTTTCTTTCACGAGACCTCC 780
Db 716 GCGAATGGGGATGAAGATGAGACGAGCAATCTACTTCTCTTACGAGACTTCC 775
QY 781 CGAGTGTGGACTCTATGAGCGCATCAAGGTCCCAAGAGTGGCGGAGTGTGCGGG 840
Db 776 CGAGCATTTGACTATACGAGCECATTAAGTCCACAGGCTGCGCGTGTGTGCGGG 835
QY 841 GACCTTGGGGAGAGACCTTTCAGCAGAGATGAGACACCTTCTGAAGCGCTGACCTG 900
Db 836 GACCTGGGGGCGGAAGACCTTCAGCAGAGATGAGACACGCTTTTGAAGCTGACCTG 895
QY 901 CTGTGCCAGGCCGAGCATGGCGGCTCCGGGTTCTCAGGCTATGCGAGGCTT 960
Db 896 CTCTGTCCAGGCCCTGAGCATGGCGGCTCCAGTGTCTCAGGATGTTCTGTGCTT 955
QY 961 CGGCCTCAGCCTGAGCGGGAACCCCATCTTTTATGGATCTTTTCTCCAGTGGGAA 1020
Db 956 CGACCTGAGCTTGGGCGAGGACTCCCATCTTTTATGGCATCTTTCTCCAGTGGGAG 1015
QY 1021 GGAGCTGCCATCTCTGCTGTGTGCTTCCGACCCCAAGACATCGGGGAGTGTGAT 1080
Db 1016 GGGGTACTATCTCTGCTGTGCTTCCGACCAACAGACATTCGGACAGTGTGAT 1075
QY 1081 GGTCCCTTAGAGACTAAACATGACTGCAACAGGAGCTGCTGTATGAGCAACAG 1140
Db 1076 GGTCCCTCAGAGAACTAAACATGACTGCAACAGAGGACTGCCGTGTGGCAATGAT 1135
QY 1141 GTGCCCCAGCCAGCCTGGAGAGTGCATCGCCAAACATGAAGCTCCAGCAGTTTGA 1200
Db 1136 GTGCCCCAGCCAGCCTGGAGAGTGCATCAACCAACATGAAGCTCCGACATTTGCG 1195
QY 1201 TCCTCAGCTCTCCTGCCAGACCGCTGTCTACCTTTATCAGACACACCTCTCATGAG 1260
Db 1196 TCATCTCTCTCCTGCTGACCGCTACTCAGCTTCTATCCGGGACCACTCATGAG 1255
QY 1261 AGGCCCGTGTCCCGGCTGACGGCGCCCTCTGCTGGTCACTACAGATACAGCCTATCTC 1320

Db 1256 AGCCAGTGTTCACAGCTGATGGCCACCCCTCTCTGCTCACTACAGATACAGCCTATCTC 1315
QY 1321 AGAGTCGTGGCCACACAGGTGACACGCTCTCAGGAAAGAAATATGACGTGCTCTACCTG 1380
Db 1316 AGAGTCGTGGCCACACAGGTGACACGCTCTCAGGAAAGAAATATGATGCTCTACCTG 1375
QY 1381 GGGACAGAGATGGACACCTCCACCGGCTGTGCGCATTTGGAGCTCAGCTCAGTCTGTTG 1440
Db 1376 GGGACAGAGATGGACACCTCCACCGAGTGTGCGATCGGATCGAGCTCAGCTCAGCTTCTT 1435
QY 1441 GAGGATCTGGCCTTGTTCACAGAACACAGCCGTTGAGAGCATGAATTTGTACCAGAT 1500
Db 1436 GAAGATCTGGCCTTATTCACAGAGCACAGCCAGTTGAGAACATGAATTTGTACCAGAC 1495
QY 1501 TGGCTCTCTGGTGGCTCCCATACTGAGTGAACAAGTGAACACCAAGCACTGTGCCGT 1560
Db 1496 TGGCTCTGTGGTCCCGTACTGAGTGAACAAGTGAACAAGTGAACAAGTGAACAAGTGA 1555
QY 1561 CTCCAGAGCTGCTGGAGTGTATCTGCGCCAGGACCCCTGTGCGCTGGAGCTTCCGG 1620
Db 1556 CTCCAGAGCTGCTCAGAGTGCATCTGCGCCAGGACCCAGTGTGCTGGAGCTTCCGG 1615
QY 1621 CTGTGATGCTGTGGCCACCGCCGAGCAGCCGCGGATGGTTCAGATATAGAGTCA 1680
Db 1616 CTGGATGAGTGTGGCCCATGCCGGGAGCACCGAGGTTGGTCCAAGACATAGAGTCA 1675
QY 1681 GCGGATGTCTCTTGTGTCCAAAGAACCTTGAGAACATCCCCCTAGTGTGTTGAAGTT 1740
Db 1676 GCAGATGTCTCTTGTGTCTTAAGAGCCTGGAGACGCTCCAGTACTGTTGAAGTT 1735
QY 1741 CCGGTGGCTACTGTGGGCCACGTGTGCTGCGCATGTTCCCAAGTGTCTGCTGGGCACTCC 1800
Db 1736 CCGGTGGCTACTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1795
QY 1801 TCTGTGTGCAACAGCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
Db 1796 TGTGTGTGCAACAGCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1855
QY 1861 GTGGTGTGCAACAGCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
Db 1856 GTGGTGTGCAACAGCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1915
QY 1921 GCGCGGTGGTGGCTTATAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
Db 1916 GCGCATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1975
QY 1981 GCGCACACGTTGTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
Db 1976 GCGCACAC---AGTGGGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2032
QY 2041 CTCACCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
Db 2033 CTGACTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2092
QY 2101 AGAGACAGGTGGGCTTATGATCTGGGGCTCACCTTCTGGGACCAACAGCTATAGTCA 2160
Db 2093 AGAGACAGGTGGGCTTGGACCTGGGGCTCCACCTTCTGGGACCAACAGCTATAGTCA 2152
QY 2161 GACCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
Db 2153 GACCT 2212
QY 2221 AGTGGTTTTGGTGGCT 2280
Db 2213 AGTGGTTTTGGTGGATTTCTACACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2272
QY 2281 ATCCGGCTACTGGGGGCTCTACCGAGTGTGATGAGACCTCCATCTA 2330
Db 2273 ATTCCGCTAACTGGGGCT 2322

[illegible]

```
Db 1436 GAAGATCTGGCCCTATTCCACAGAGCCACAGCCAGTTGAGAACATGAATTTGTACCACAGC 1495
Qy 1501 TGGCTCTCTGTGGGCTCCCATCTAGAGTGACAAAGTGAACACAGCAACTGTGGCCGT 1560
Db 1496 TGGCTCTCTGTGGGCTCCCATCTAGAGTGACAAAGTGAACACCAACTGTGGCCGT 1555
Qy 1561 CTCAGAGCTGTCGAGGTATCTCTGGCCAGGACCCGCTGTGCGCCCTGGAGCTTCCGG 1620
Db 1556 CTCAGAGCTGTCAGAGTGCATCTCTGGCCAGGACCCAGTCTGTGCTGGAGCTTCCGG 1615
Qy 1621 CTTGATGCTGTGTGGCCACGCGCGGAGCACCAGCCGGGATGTTCAAGATATAGAGTCA 1680
Db 1616 CTGGATGAGTGTGTGGCCATGCGCGGGAGCACCGAGGGTGTGTCACAGACATAGAGTCA 1675
Qy 1681 GCGGATGCTCTCTCTTTGTGTCACAAAGAACCTGGAGAACATCCCTAGTGTGTTGAAGTT 1740
Db 1676 GCAGATGCTCTCTCTTTGTGTCCTAAAGAGCTGGAGAGCTCCAGTAGTGTGAAGTT 1735
Qy 1741 CCGGTGGCTACTGTGGCCACGCTGTGTCCTGCGCATGTTCCGCCAGTTCCTGCCCTGGGCATCC 1800
Db 1736 CCGGTGGCTACTGTCGATGCTGTGTCCTGCGCATGTTCCGCCAGTTCCTGCCCTGGGCATCC 1795
Qy 1801 TGTGTGTGACACAGCCAGTGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1860
Db 1796 TGTGTGTGACACAGCCAGTGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1855
Qy 1861 GTGGTGTGACCCAGGCGGCTATGCTTGGAGTGTGAGGAGGTGGAGCC 1920
Db 1856 GTGGTGTGACCCAGGCGGCTATGCTTGGAGTGTGAGGAGGTGGAGCC 1915
Qy 1921 GCGCGGTGGTGTGCTTATAGTGTGGTGTGGGCGACGACGCGGACCCCAACCGG 1980
Db 1916 GCGCATGTGTAGCAGCTTACAGCTTGGTATGGGCGAGCAGGAGATGCTCCGAGCCGG 1975
Qy 1981 GCGCACCGTGTGGGCGCTGGATGGTGGCTTTCCTCTGGGTGTTCTTGACGATCC 2040
Db 1976 GCGCACAC---AGTGGGCGGAGTGGTGGCTTTCCTCTGGGTGTTCTTGACGATCC 2032
Qy 2041 CTCACCTCTCTCTCTGATGTCGCTCAGCAGCTGCGCGACAGAGGAGCTTCTAGCT 2100
Db 2033 CTGACTCTCATCTGATGTCGCGCTCAGCAGGAGCGGACAGAGGAACTTCTGGCT 2092
Qy 2101 AGAGCAAGGTGGGCTTAGATCTGGGGCTCCACCTCTGAGACCAAGCTATAGTCAG 2160
Db 2093 AGAGCAAGGTGGGCTTGACCTGGGGCTCCACCTCTGAGACCAAGCTATAGTCAG 2152
Qy 2161 GACCTCTCTCTCTCTGCTGAGATGAAGCTGCGCTGCGCTGCGCTGCGCTGAGCGGCG 2220
Db 2153 GACCTCTCTCTCTCTCTGAGATGAGCGGTGCGCTGCGCTGCGCTGAGCGGCG 2212
Qy 2221 AGTGGTGTGGTGGCTTCCCTCCACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
Db 2213 AGTGGTGTGGTGGATCTCACCACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2272
Qy 2281 ATCCGGCTCAGTGGGCGCTCTAGCCAGCTGTGATGAGACCTCCATCTA 2330
Db 2273 ATTCGGCTTAAGTGGGCTCTCTAGCCACATGTGATGAACATCTCATCTA 2322
```

RESULT 15

US-09-160-762-3

```
; Sequence 3, Application US/09160762A
; GENERAL INFORMATION:
; APPLICANT: David Michaelovich
; APPLICANT: Trudy Rachel Doe
; APPLICANT: Philip David Hayes
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP-30036
; CURRENT APPLICATION NUMBER: US/09/160,762A
; CURRENT FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows version 3.0
```

```
; SEQ ID NO 3
; LENGTH: 2894
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
; US-09-160-762-3

Query Match 69.8%; Score 1626; DB 16; Length 2894;
Best Local Similarity 83.2%; Pred. No. 0;
Matches 1938; Conservative 0; Mismatches 275; Indels 117; Gaps 3;

Qy 1 ATGCTTTGCCAGGCGCCAGCGGCGCCCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
Db 110 ATGCGCGCGCTCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 169
Qy 61 CGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db 170 CC-----GCTACTGCTGCTGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 214
Qy 121 GTCCCGCGCTCAGTGGCCAGAACCTCGCTGCCCATCTCCGAGGCTGACTCCTATCTCACC 180
Db 215 GTCCCGCGCTCGGTGCCAGAACCTCGCTTCCAATCTCTGAGGCTGACTCCTGCTCACC 274
Qy 181 CGGTTTGCAGCGCTCTCATAGTACATTTACTCTGCTCTCTCTGCTGCTGCTGCTGCTGCT 240
Db 275 CGGTTGCGAGTCCCTCACACATACAAATTTACTCTGTTCTCTCTGCTGCTGCTGCTGCTG 334
Qy 241 ACACCTTACGTCGGTGCAGCGGATAGCATCTTCGCTTTAAACCTCCCGCTTCTCTGCGGAA 300
Db 335 ACACCTTATGTTGGCGCCCGGACACCATCTTTCGCTTTATCTCTGCTGCTGCTGCTGCTG 394
Qy 301 AGACCCCGAAGATCGACTGGATGATGCTGAGACTCACAGACAGAACTGCAGGAAGAAA 360
Db 395 AGACCCCGCAGGATGACTGGATGCTTCTGAGGCTCACAGACAGAACTGTAGGAAGAAA 454
Qy 361 GSCAAGAAAGAGGACCAATGTCACAATTTATCCAGATTCGCCCAATGTCATGCTCT 420
Db 455 GSCAAGAAAG----- 466
Qy 421 CACCTCTCACGTCGGGCACCTTTCGCTTTTGTGATCCGAGTGGGGGTATTGATGTGTC 480
Db 467 -----GATGTGTC 475
Qy 481 AGTTTCCAGCAGGTGAAAGACTTGAGAGCGCGCGCGGGAATGTCCTTTTGTGAGCCAGCT 540
Db 476 AGGTTCCAGCAGGTGAAAGACTTGAGAGTGGCGCGGGAATGTCCTTTTGTGAGCCAGCT 535
Qy 541 CAACGCTCAGCAGCTGTAATGCTGGGCGCTCTCTACACGCGCACTGTGAAGAACTTC 600
Db 536 CAGCGGTGACAGCTGTAATGCTGGGGGGTCTCTATGCTGCCACTGTGAAGAACTAC 595
Qy 601 CTGGGACTTGAGCCCATCATCTCCGAGCTGTGGTTCGAGCTGAGACTGGATTCGAACA 660
Db 596 CTGGGACGAGCAATTTATCAGAGCAGTGGGTGCTGCGGAGGACTGGATTCGAGCA 655
Qy 661 GAGACCTGTCTATCTGGCTTAATGCTCAGCGCTTGTGCGAGCTATGCTGCTGAGCCCA 720
Db 656 GATACCTTGCCTTCTGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 715
Qy 721 GCTGAGTGGGGGATGAAGATGAGAGATGAATCTTTTCTTTTCTTCTACGAGACTTCC 780
Db 716 GCCGATGGGGGATGAAGATGAGAGACGACGAAATCTACTTCTTTTACGAGACTTCC 775
Qy 781 CGAGTGTGGACTCTCTATGAGCGCATCAAGGTCCCAAGAGTGGCGGAGTGTGTGCGGG 840
Db 776 CGAGCATTTGACTCATACGAGCGCATTAAGTCCCAAGGCTGGCGCGTGTGTGCGGGG 835
Qy 841 GACCTTGGGGGAGGAAAGCCCTTACAGAGATGAGACGCTTTCTGAAGGCTGACCTG 900
Db 836 GACCTCGGGGCGGGAAGACCCCTCCAGCAGATGAGACGCTTTTGTGAAGCTGACCTG 895
Qy 901 CTGTGCCAGGCGCGGAGCATGCGCGGCTCCGGGCTTCTCAGGCTATGTCAGAGCTT 960
Db 896 CTCTGTCCAGGCGCTGAGCATGCGCGGCTCCAGTGTCTCAGAGATGTTGCTGTGCTT 955
```

Qy	961	CGGCTCAGCCTGGAGCGGGAACCCCCATCTTTTATGGGATCTTTTCCTCCCAGTGGGAA	1020
Db	956	CGACTGAGCTTGGGCGAGGACTCCCATCTTTTATGGCATCTTTTCTTCCCAGTGGGAG	1015
Qy	1021	GGAGTGCCATCTCTGTGTGTGCCCTTCGGACCCTCAAGACATCCGGGCAGTGCCTGAAT	1080
Db	1016	GGGCTACTATCTCTGTCTGTGTGCCCTTCGGACCACAAGACATTCGACAGTGTCTGAAT	1075
Qy	1081	GGTCCCCTTTAGAGACTAAAACATGACTGCAACAGGGGACTGCTGTCAATGGACAACGAG	1140
Db	1076	GGTCCCCTCAGAGAACATAAACATGACTGCAACAGAGGACTGCTGTGCTGGACAATGAT	1135
Qy	1141	GTGCCCCAGCCAGACCTTGAGAGTGCATCGCCACAACATGAAGCTCCAGCAGTTTGGGA	1200
Db	1136	GTGCCCCAGCCAGACCTTGAGAGTGCATCACCAACAACATGAAGCTCCGGCAGCTTTGGC	1195
Qy	1201	TCTCTACTCTCCCTGCCAGACCGCGTGCTCACCTTTATCAGAGACCAACCCTCTCATGGAC	1260
Db	1196	TCATCTCTCTCCCTGACCGGTACTIONCTTATCCGGGACCACCACCATCATGGAC	1255
Qy	1261	AGGCCCTGTTCGGGGTGACGGCCGCCCTGCTGTGTACTACAGATACAGCCTATCTC	1320
Db	1256	AGGCCAGTGTTCAGCTGTATGGCCACCCCTGCTGTGTACTACAGATACAGCCTATCTC	1315
Qy	1321	AGAGTCGTGCCACAGGGTGACCAGCCTCTCAGGGAAGAAATATGACGTGCTCTACCTG	1380
Db	1316	AGAGTCGTGCCACAGGGTGACAGCCTCTCAGGGAAGAGTATGATGTCTCTACCTG	1375
Qy	1381	GGGACAGAGATGGACACCTCCACCGGGCTGTGGCATTGGAGCTCAGCTCAGTGTCTTG	1440
Db	1376	GGGACAGAGATGGACACCTCCACCGAGCAGTGGGATCGGAGCTCAGCTCAGCTGTCTT	1435
Qy	1441	GAGATCTGCCCTGTFTCCAGAACACACAGCCGTTGAGAGCATGAAATTTGTACCAAGAT	1500
Db	1436	GAAGATCGCCCTATTTCACAGAGCCACAGCCAGTTGAGAACATGAAATTTGTACCACAGC	1495
Qy	1501	TGGCTCTGTGGGCTCCCATACTGAGGTGACACAAAGTGAAACACCACCACTGTGGCCGT	1560
Db	1496	TGGCTCTGTGGTCCCTGCTACTGAGGTGACACAAAGTGAATACAACCACTGTGGCCGT	1555
Qy	1561	CTCCAGAGCTGCTCGGAGTGTATCTTGGCCACAGACCCCTGTGGCCTCGAGCTTCCGG	1620
Db	1556	CTCCAGAGCTGCTCAGAGTCATCTTGGCCACAGGACCCAGTCTGTGCTCGAGCTTCCGG	1615
Qy	1621	CTTGATGCTGTGTGGCCACAGCCGGGGAGACACCGGGATGTTCAAGATATAGAGTCA	1680
Db	1616	CTGATGAGTGTGTGGCCCATGCGGGGGAGCACCGAGGTTGGTCTCAAGACATAGAGTCA	1675
Qy	1681	CGGGATGCTCTCTTTGTGTCAAAGAAGACCTGGAGAACATCCGTAGTGTGTGAGTT	1740
Db	1676	GCAGATGCTCTCTTTGTGTCTTAAAGAGCCTGGAGAACGTCCAGTAGTGTGTGAGTT	1735
Qy	1741	CCGGTGGCTACTGTGGCCACGTGGTGCTTGGCCATGTTTCCCCAGTGTCTGCTCGGCATCC	1800
Db	1736	CCGTGGCTACAGCTGGGCATGTGGTCTTGCCATGTTCTCAAAGCTCAGCATGGGCATCC	1795
Qy	1801	TGTGTGTGGACAGACCCAGTGGATGACTGCGCTCACTCCCGGAGGGATGAGCATAGAG	1860
Db	1796	TGTGTGTGGACAGCCAGTGGAGTACTGTCACTACCCCGCGGGATGGAGCTGAG	1855
Qy	1861	GTGGTGTGACCCAGSGGCCATGGGGGCTTATGCTTGGCAGTCTCAGGAGGTTGGAGCC	1920
Db	1856	GTGGTGTGACCCAGSGGCCATGGGGGCTTATGCTTGTGAATGTCAAGAGGTTGGGGCA	1915
Qy	1921	GCCCGCTGTGGTGTCTTAGCTTGGTGTGGGGACGCCAGCGGGGACCCCTCAAAACCGG	1980
Db	1916	GCCCATGTGTAGCAGCTTACAGCTTGTGTATGGGACGCCAGCATGCTCCGAGCCGG	1975
Qy	1981	GCCACACCGTGTGTGGGGCTGGATGGTTGGCTTCTCTCTGGGTGTCTTGTTCAGCATCC	2040
Db	1976	GCCACAC --- AGTGGGGGGGAGACTGGCTGGCTTCTTCTTGGGGATTCFCGAGCATCC	2032

Search completed: September 26, 2003, 10:54:48
Job time : 6800.65 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 26, 2003, 02:11:50 ; Search time 360.737 Seconds
(without alignments)
8552.073 Million cell updates/sec

Title: US-09-284-180A-2

Perfect score: 2331

Sequence: 1 atgttgccgggcccggagc.....gtgatgagacctcatctaa 2331

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1152032 seqs, 66174393 residues

Total number of hits satisfying chosen parameters: 2304064

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_NA_New.*

- 1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
- 2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
- 3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
- 4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
- 5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
- 6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
- 7: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1271.2	54.5	2082	1	PCT-US03-27411-47
2	248.6	10.7	1914	1	PCT-US03-09929-1
3	248.6	10.7	1914	1	PCT-US03-09929-7
4	248.6	10.7	1914	1	PCT-US03-09929-9
5	247.8	10.6	3112	1	PCT-US03-09929-5
6	246	10.6	2739	1	PCT-US03-09929-11
7	223.2	9.6	2049	1	PCT-US03-09929-3
8	215.2	9.2	2589	1	PCT-US02-36071A-168
9	215.2	9.2	4157	1	PCT-US02-36071A-167
10	205.6	8.8	2586	1	PCT-US02-36071A-165
11	205.6	8.8	2769	1	PCT-US02-36071A-164
12	155.6	6.7	2705	1	PCT-US03-27411-44
13	145	6.2	725	1	PCT-US03-27411-40
14	62	2.7	1327	1	PCT-US03-09929-21
15	62	2.7	1327	1	PCT-US03-17412-21
16	62	2.7	1438	1	PCT-US03-09929-25
17	62	2.7	1438	1	PCT-US03-17412-25
18	62	2.7	1492	1	PCT-US03-09929-23
19	62	2.7	1492	1	PCT-US03-17412-23
20	62	2.7	1878	1	PCT-US03-09929-33
21	62	2.7	1878	1	PCT-US03-17412-33
22	62	2.7	1908	1	PCT-US03-09929-35
23	62	2.7	1908	1	PCT-US03-17412-35
24	62	2.7	1921	1	PCT-US03-09929-15
25	62	2.7	1921	1	PCT-US03-17412-15
26	62	2.7	1948	1	PCT-US03-09929-49

27 62 2.7 1948 1 PCT-US03-17412-49 Sequence 49, Appl
28 62 2.7 2113 1 PCT-US03-09929-55 Sequence 55, Appl
29 62 2.7 2113 1 PCT-US03-17412-55 Sequence 55, Appl
30 62 2.7 2583 1 PCT-US03-09929-51 Sequence 51, Appl
31 62 2.7 2583 1 PCT-US03-17412-51 Sequence 51, Appl
32 62 2.7 2634 1 PCT-US03-09929-53 Sequence 53, Appl
33 62 2.7 2634 1 PCT-US03-17412-53 Sequence 53, Appl
34 62 2.7 2944 1 PCT-US03-09929-29 Sequence 29, Appl
35 62 2.7 2944 1 PCT-US03-17412-29 Sequence 29, Appl
36 62 2.7 2995 1 PCT-US03-09929-19 Sequence 19, Appl
37 62 2.7 2995 1 PCT-US03-17412-19 Sequence 19, Appl
38 62 2.7 3055 1 PCT-US03-09929-27 Sequence 27, Appl
39 62 2.7 3055 1 PCT-US03-17412-27 Sequence 27, Appl
40 62 2.7 3106 1 PCT-US03-09929-17 Sequence 17, Appl
41 62 2.7 3106 1 PCT-US03-17412-17 Sequence 17, Appl
42 62 2.7 3165 1 PCT-US03-09929-47 Sequence 47, Appl
43 62 2.7 3165 1 PCT-US03-17412-47 Sequence 47, Appl
44 62 2.7 3498 1 PCT-US03-09929-31 Sequence 31, Appl
45 62 2.7 3498 1 PCT-US03-17412-31 Sequence 31, Appl

ALIGNMENTS

RESULT 1
PCT-US03-27411-47
; Sequence 47, Application PC/TUS0327411
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; Lee, Soo Yeun;
; APPLICANT: ELLIOTT, Vicki S.; HAFALIA, April J.A.;
; APPLICANT: BURFORD, Neil; SPRAGUE, William W.;
; APPLICANT: GRIFFIN, Jennifer A.; YANG, Yonghong G.;
; APPLICANT: CHAWLA, Narinder K.; BAUGHN, Mariah R.;
; APPLICANT: BECHA, Shanya D.; KHARE, Reena;
; APPLICANT: THORNTON, Michael B.; MASON, Patricia M.;
; APPLICANT: GIETZEN, Kimberly J.; ISON, Craig H.;
; APPLICANT: MARQUIS, Joseph P.; SWARNAKAR, Anita;
; APPLICANT: RANKUMAR, Jayalaxmi; JIN, Pei;
; APPLICANT: RICHARDSON, Thomas W.; TRAN, Uyen K.
; TITLE OF INVENTION: NEUROTRANSMISSION-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-1557 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/27411
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 60/408,383
; PRIOR FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 60/408,781
; PRIOR FILING DATE: 2002-09-05
; PRIOR APPLICATION NUMBER: US 60/414,221
; PRIOR FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/426,483
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 60/431,566
; PRIOR FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: US 60/434,317
; PRIOR FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US 60/437,763
; PRIOR FILING DATE: 2003-01-03
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PERL Program
; SEQ ID NO 47
; LENGTH: 2082
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 7523644CBI
PCT-US03-27411-47

Query Match 54.5%; Score 1271.2; DB 1; Length 2082;
Best Local Similarity 82.4%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 213; Indels 114; Gaps 2;

QY 1 ATGCTTGCCAGGCGCGGAGCGGCGCGCGCGCGCGCGCGCGCTCCGCTCTTCCCTTC 60

; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 7
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1914)
PCT-US03-09929-7

Query Match 10.7%; Score 248.6; DB 1; Length 1914;
Best Local Similarity 52.1%; Pred. No. 5.3e-58;
Matches 704; Conservative 0; Mismatches 629; Indels 18; Gaps 6;

```
QY 308 GAAGGATGACTGGATGGTACCTTGAGACTCACAGACAGAACTCCAGGAGAGAAGGAAG- 366
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 200 GAGCGATCTCTGGAGGCCCCCGTGGAGGAGNAGACTGAGTGATCCAGAGAAGGAAGA 259

QY 367 --AAAGAGACGAATGTCAAAATTTATCCAGATTCGCCCAATGTGCAATGCTCTCACC 424
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 260 ACAACAGACCGAGTCTTCAACTTCATCGCTTCTCGAGGCCCTACAATGCTCCACC 319

QY 425 TCCTCAGTGGCGACCTTCGCTTTTGATCCGAAAGTGGGGTTATGATGTTGTCAGTT 484
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 320 TGTACGCTGTGGCACTAGCGCTTCCAGGCCAAAGTGCACCTACGCTCAACATGCTCACCT 379

QY 485 TCAC---GCAGGTTGAAGACTTTGAGAGCGCGCGGGGAAATGTCCTTTTGAGCCAGCTC 541
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 380 TCACTTTGGAGCATGAGAGATTTGAAGATGGGAGGCAAGTGTCTCTATGACCCAGCTA 439

QY 542 AAGGTCAGCAGCTGTAATGGCTGGGGGCTCTCTACACCGCCACTGTGAAGAATCTCC 601
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 440 AGGGCCATGCTGGCCCTTCTTGTGGATGGTGAAGTGTACTTCTTCAGGGAGCGGGCAG 499

QY 602 TGGGGACTGAGCCCATCATCTCCGAGCTGTGGTCCAGCTGAGGACTGGATTCGAACAG 561
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 500 TGGGCACGGAACCAATATCTCTGCGTAAACATGGGGC---CCCACCACTCCATGAAGACAG 556

QY 662 AGACCTTGTATCTCTGGCTTAATGCTCCAGCTCTTGTCCAGCTGATGGTCTGAGCCAG 721
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 557 AGTACCTGGCTTTGGCTCAGCAACCTCACTTTGTAGGCTGCTGCTATGCTACCTGAGA 616

QY 722 CTGAGTGGGGGATGAAGATGGAGCAGATGAATCTTTTTCACGGAGACCTCC 781
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 617 GTGTGGCAGCTTACGGGGGACGACGACAAGTGTACTTCTTTCAGGGAGCGGGCAG 676

QY 782 GAGTGTGACCTCTATGAGCCATCAAGTCCCAAGAGTGGCCCGAGTGTGTGGGGGG 841
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 677 TGGAGTCCGACTGCTATGCCAGCAGGTGGTGGCTGCTGTGGCCCGTGTCTCAAGGGCG 736

QY 842 ACCTTGGGGGAGGAGACCTTTCAGCAGAGATGGACGAGTCTTCTGAAGGCTGACCTGC 901
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 737 ATATGGGGGCGCAGGACCTTCAGAGGAAGTGGACCACTTCTCTGAAGGCGGGCTGG 796

QY 902 TGTGCCAGGGCGGAGCAGATGGCGGGCTCCGGGTTCTGCAGGCTATGGCAGAGCTTC 961
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 797 CATGCTCTGCCCGAAGTGGCAGCTTCTTCAACCAAGCTGCGAGGGGATG---CACACCC 853

QY 962 GGCCTCAGCTGGAGGGGAACCCCATCTTTATGGGATCTTTCTCCAGTGGGAG 1021
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 854 TCGAGGACCTCTCTGGCAACAACACCATCTTTTGGGGTTTTTCAAGCAGAGTGGGGTG 913

QY 1022 GAGCTGCCATCTCTGTGTGTGCTTCCGACCCCAAGACATCCGGGCGAGTGTGAATG 1081
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 914 ACATGTACCTGTGGCCATCTGTGATACCATCTTGAAGAGATCCAGGGGTGTTGAGG 973

QY 1082 GTCCCTTTAGAGCTAAACATGACTGCAACAGGGGAGTGCCTGTCTATGGACACAGAG 1141
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 974 GCGCCCTAAGGATACCATGAGGAAGCCAGAAAGTGGGACCGCTACACTGACCT---G 1030

QY 1142 TCGCCAGCCAGCTGGAGAGTGCATCGCCAAACATGAAGCTCCAGCAGTTTGGAT 1201
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 4

PCT-US03-09929-9

; Sequence 9, Application PC/TUS0309929

; GENERAL INFORMATION:

; APPLICANT: Curagen Corporation, et al.

; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES,

; FILE REFERENCE: 21402-573B-061

; CURRENT APPLICATION NUMBER: PCT/US03/09929

; CURRENT FILING DATE: 2003-04-01

; PRIOR APPLICATION NUMBER: 60/368,996

; PRIOR FILING DATE: 2002-04-01

; PRIOR APPLICATION NUMBER: 60/369,980

; PRIOR FILING DATE: 2002-04-04

; PRIOR APPLICATION NUMBER: 60/370,381

; PRIOR FILING DATE: 2002-04-05

; PRIOR APPLICATION NUMBER: 60/370,969

; PRIOR FILING DATE: 2002-04-08

; PRIOR APPLICATION NUMBER: 60/371,002

; PRIOR FILING DATE: 2002-04-09

; PRIOR APPLICATION NUMBER: 60/372,002

; PRIOR FILING DATE: 2002-04-12

; PRIOR APPLICATION NUMBER: 60/384,297

; PRIOR FILING DATE: 2002-05-30

; PRIOR APPLICATION NUMBER: 60/386,816

; PRIOR FILING DATE: 2002-06-07

; PRIOR APPLICATION NUMBER: 60/389,123

; PRIOR FILING DATE: 2002-06-13

; PRIOR APPLICATION NUMBER: 60/402,207

; PRIOR FILING DATE: 2002-08-09

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 179

; SOFTWARE: CuraseqList version 0.1

; SEQ ID NO 9

; LENGTH: 1914

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(1914)
PCT-US03-09929-9

Query Match.	10.7%	Score 248.6;	DB 1;	Length 1914;
Best Local Similarity	52.1%	Pred. No. 5.3e-59;		
Matches 704; Conservative	0;	Mismatches 629;	Indels 18;	Gaps 6;
QY	308	GAAGGATCGACTCGATGGTACCTGAGACTCAGACAGAACTCAGAGAAAGGCAAGA	367	
DB				
QY	200	GACGATCTCTGGGAGGCCCGTGGAGAAGAGACTGAGTGTATCCGAAAGGGAAGA	259	
DB				
QY	368	A---GAGGACGATGTCAAAATTTATCCAGATTTCTGCGCATTTGTCATGTCCTCTCACC	424	
DB				
QY	260	ACAGCCAGACCGAGTGTCTCAACTTCTATCCGCTTCTGCGACCTCAATGCTCCCAACC	319	
DB				
QY	425	TCTCTACGTCGCGCACCTTCGCTTTTGATCCGAAGTCGGGGTTATTGATGTGTCCAGTT	484	
DB				
QY	320	TGTACGTCGTGTGCACCTACGCTTCCAGCCCAAGTCACACTGCTCAACATGCTCACT	379	
DB				
QY	485	TCCA---GCAGTTGAAAGACTTGTAGAGCGCGCGGGAATGTCCTTTTGAGCCAGCTC	541	
DB				
QY	380	TCATTGGAGCATGGAGATTGAGATGGGAAGGCAAGTGTCCCTATGACCCAGCTA	439	
DB				
QY	542	AACGGTCAGCAGTGTAAATGGCTGGGGCGTCTCTATACCGCACTGTGAAGAACTTCC	601	
DB				
QY	440	AGGGCATGCTGGCTTCTTGTGGATGGTGAGCTGTACTCGGCCACACTCAACAACCTTC	499	
DB				
QY	602	TGGGGACTGAGCCCATCATCTCCCGAGCTGTGGTCGAGCTGAGGACTGATTCGACAG	661	
DB				
QY	500	TGGGCAGGAACCAATATCTCGGTAAACATGGGGC---CCCACCACTCCATGAAGACG	556	
DB				
QY	662	AGACCTGTGTCATCTGGCTTAAATGCTCCAGCCTTTGTGCGAGCTATGGTCTGAGCCAG	721	
DB				
QY	557	AGTACCTGGCTTTGGCTCAACGAACCTCACATTGTAGGCTCTGCCATATGTACTCTAGA	616	
DB				
QY	722	CTGAGTGGGGGATGAAGATGGAGAGATGAAATCTTTTCTTCTCACGGAGACCTCC	781	
DB				
QY	617	GTGTGGGCAGCTTTCACGGGGGAGCAGACAAGGTCTACTTCTTCTTCAGGAGCGGGCAG	676	
DB				
QY	782	GAGTGTGGACTCCTATGAGCGCATCAAGCTCCAGAGTGGCCGAGTGTGTCGGGG	841	
DB				
QY	677	TGAGTCCGACTGCTATGCCGAGAGGTGTGGCTGTGTGGCCGGTGTCTGCAAGGGCG	736	
DB				
QY	842	ACCTTGGGGCAGGAAGACCTTTCAGCAGAGATGGAGAGCTTTCTGAAGGCTGACCTGC	901	
DB				
QY	737	ATATGGGGGCGCAGGACCTTCAGAGGAAGTGGACCAGTTCCTTGAAGCGCGGTGG	796	
DB				
QY	902	TGTGCCAGGGCCGAGCATGGCCGGGCTTCGGGGTTCTGAGGCTATGGCAGAGTTC	961	
DB				
QY	797	CATGCTCTGCCCGAATGTCAGCTCTACTTCAACACAGCTGCAGGCGATG---CACACC	853	
DB				
QY	962	GGCTCAGCTTGGAGGGGAACCCCATCTTTATGGGATCTTTCTCCCGAGTGGGAAG	1021	
DB				
QY	854	TGCAGACACTCTGTGCACACACCACTTCTTTGGGGTTTTTCAAGCAGAGTGGGGT	913	
DB				
QY	1022	GAGCTGCCAFTCTGTGTGTGCTTCCGACCCCAAGACATCCGGGCACTGCTGAATG	1081	
DB				
QY	914	ACATGTACCTGTGGCCATCTGTGAGTACCAGTTGGAAGATCCACGCGGTGTTTGAGG	973	
DB				
QY	1082	GTCCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGTCTGTGTATGGACAGCAGG	1141	
DB				
QY	974	GCCCTTATAAGAGTACCATGTAGGAAGCCAGAGTGGGACCGCTACACTGACCT---G	1030	
DB				
QY	1142	TGCCCCAGCCACACTGGAGTGTGCATCCCAACACATGAAGCTCCAGCACTTGGAT	1201	
DB				
QY	1031	TACCCAGCCTCGGCCTGGCTGTGTGATTAACAACTGGCATTCGGCGCCCGGTACACCA	1090	
DB				
QY	1202	CTTCACTCTCCCTGCCAGACCGGTGCTACCTTTTATCAGAGACCACTCTCTCATGGACA		

Db 462 ACAATGCTCCCACTGTACGTCTGTGGCACCTACGCTTCCAGCCCAAGTGCACCTACG 521
Qy 470 TTGATGTGTCCAGTTTCCCA---GCAGGTTGAAAGACTTGAGAGCGCGCGGGGAAATGTC 526
Db 522 TCACATGCTCACCTTCCATTTGGAGATGGAGAGTTTGAAGATGGGAAGGCAAGTGTGTC 581
Qy 527 CTTTGTGACGAGCTCAACGGTCAAGAGCTGAGAGCTGTAATGGCTGGGGGGTCTCTACACGCCCA 586
Db 582 CCTATGACCCAGCTAAGGGGCATGCTGGCCCTTCTGTGGATGGTGTGAGCTGTACTCGGCCA 641
Qy 587 CTGTGAAGAACTTCTCGGGGACTGAGCCCATCATCTCCGAGCTGTGGGTGCGAGCTAGG 646
Db 642 CACTCAACAATCTCTGGGACGGAACCCATTAATCTCGCTGAACATGGGG---CCACAC 698
Qy 647 ACTGGATTCAACAGACACTTGTCACTCGCTGGCTTAATGCTCCAGCCCTTGTCTCCAGCTA 706
Db 699 ACTCCATGAACAGAGACTACTGGCCCTTTGGCTCAACGAACCTCACTTTGTAGGCTGTG 758
Qy 707 TGGTCTGAGCCCAAGCTGAGTGGGGGATGAAGATGGAGACGATGAATCTTTTCTTCT 766
Db 759 CCTATGTACCTGAGAGTGTGGGACGCTTCAACGGGGGACGACGAAGTCTACTTCTCT 818
Qy 767 TCAGGAGACCTCCCGAGTGTGGACTTCTATGAGCCCATCAAGGTCCCAAGATGGGCC 826
Db 819 TCAGGAGCGGGAGTGGAGTCCGCTGCTATGCCGAGCAGGTGGTGGCTCGTGTGGGCC 878
Qy 827 GAGTGTGTGGGGGACCTTGGGGCAGGAAGCCCTTACGACAGATGGACGAGTTTC 886
Db 879 GTGTCTCAAGGGGATATGGGGGCGCAGCGACCTTGCAGGAAGTGGACACGTTCC 938
Qy 887 TGAAGGTGACCTGCTGTGCCAGGGCCGAGCATGCGCGGCTCCGGGTTCTCGAGG 946
Db 939 TGAAGGCGGCTGGCATGCTCTGCCCGAACTGGCAGCTCTACTTCAACAGCTGCAGG 998
Qy 947 CTATGGCAGAGTTCGGCTCAGCTGGAGCGGGAACCCCATCTTTTATGGGATCTTT 1006
Db 999 CGATG---CACACCTGCAGACACCTCTCTGGCAACAACACCACTCTTTGGGGTTTTC 1055
Qy 1007 CCTCCAGTGGGAAGAGCTGCCATCTCTGTGTGTGTGCTCCGACCCCAACACATCC 1066
Db 1056 AAGCACAGTGGGTGACATGTACTCTGTGGCCATCTGTGAGTACCAGTTGGAAGATCC 1115
Qy 1067 GGGCAGTGTGAATGTGTCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTCGCTG 1126
Db 1116 AGCGGTGTGTGAGGGCCCTATTAAGGAGTACCATGAGGAAGCCAGAAGTGGACCGCT 1175
Qy 1127 TCATGGACAGGAGTGGCCAGCCAGACCTGGAGAGTGCATGCCCAACAATGAAGC 1186
Db 1176 AACTG---ACCCGTGTACCAGCCCTCGGCTGCTGCTGCAATTAACAACCTGGCATCGC 1232
Qy 1187 TCCAGCAGTTTGGATCCTACTCTCCCTGCGACAGCCGCTGCTCACTTTATCAGAGCC 1246
Db 1233 GCCAGGCTACACAGCTCCTCGGAGTACCCGACACATCTCACTTCTCAAGAGC 1292
Qy 1247 ACCCTCTATGAGCAGCCCGTGTCCCGCTGACGGCCGCCCTGCTGTCTACTACAG 1306
Db 1293 ACCCGCTGATGGAGGAGCAGTGGGGCTCGTGGAGCGCCGCTGCTGTGGAAGAAG 1352
Qy 1307 ATACAGCCTATCTCAGAGTGTGGCCCAAGGGTGCACGCTCTCAGGGAAGAATATG 1366
Db 1353 GCACCAACTTACCCACTGTGTGGCCGACCGGGTTACAGGACTGTGATGGAGCCACCTATA 1412
Qy 1367 ACCTGCTTACTCTGGGACAGAGGATGGACACCTCCACCGGCTGTGGCATTGGAGCTC 1426
Db 1413 CAGTGTCTTCAATGGGACAGGTCAAGCATGGTGTCTCAGGCTGTGAGCTGGGGCCT 1472
Qy 1427 AGCTCAGTGTCTGGAGGATCTGGCCCTTGT---CCGAAACCAACAGCCGTTGAGAGCA 1483
Db 1473 GGGTTCACCTGATTGAGGAGCTGCAGCTGTTTGACCAAGGAGCCCATGAGAAGCTGTGTC 1532
Qy 1484 TGAATTTGATACCAGATTTGCTCTGTGGTGGCTCCCATACTAGGTTGACACAAGTGAACA 1543

Db 1533 TATCTAGTCGACAGAAGCTGCTCTTTTCCGGCTCCCGCTCTCAGCTGGTGCAGCTGCCCG 1592
Qy 1544 CCAGCAACTGTGGCCGTCTCAGAGCTGCTCGGAGTGTATCTGCGCCAGACACCCGTGT 1603
Db 1593 TGGCCGACTGCATGAAGTATCGCTCTGTGCAGACTGTGCTCGCCGGGAGCCCTATT 1652
Qy 1604 CGCGCTGGAGTTCCGGCTTGATGCTTGTGTGGCCACAGCGGGCGA 1649
Db 1653 GCGCTGGAGCGTCAACACCAACCGCGCTGTGTGGCTGGGTGGCCA 1698

RESULT 7

PCT-US03-09929-3
; Sequence 3, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: CutaGen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND MET
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuroSeqList version 0.1
; SEQ ID NO 3
; LENGTH: 2049
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1491)
PCT-US03-09929-3

Query Match 9.6%; Score 223.2; DB 1; Length 2049;
Best Local Similarity 52.6%; Pred. No. 5.5e-51;
Matches 608; Conservative 0; Mismatches 533; Indels 15; Gaps 5;
Qy 308 GAAGACTCGACTGATGCTACCTGAGACTCACACAGAACTGCAGGAAGAAAGCAAG- 366
Db 260 GAGCGATCTCTGGGAGGCCCCCTGGAGGAGAGACTGAGTGTATCCAGAAAGGAGA 319
Qy 367 --AAAGAGGAGATGCAAAATTTATCCAGATTCGCGCAATTTGCAATGCCCTCTACC 424
Db 320 ACAACAGAGCGAGTGTCAACTTCATCCGCTTCCTGCAGCCCTACAAATGCCCTCCACC 379
Qy 425 TCCTCAGCTGGGCACTTCGCTTTTGTATCCGAAGTGGGGTTATTGATGTGCTCAGTT 484
Db 380 TGTACGTCTGTGGCACCCTACGCTTCCAGGCCAAAGTGACCTACGCAACATGCTCACCT 439
Qy 485 TCCA---GCAGTTGAAAGACTTCAGAGCGCGCGGGGAAATGCTCTTTTTCAGCCAGCTC 541
Db 440 TCACCTTTGGAGCATGGAGAGTTTGAAGATGGGAAGGCAAGTGTCCCTATGACCCAGCTA 499
Qy 542 AACGGTCAGCAGCTGTAAATGGCTGGGGCGCTCTCTACACCGCCACTGTGGAAGAACTTCC 601

```
Db 500 AGGGCCATGCTGGCCCTTCTTGAGATGTGAGCTGTACTCGGCCACACTCAACAACCTCC 559
Qy 602 TGGGACGTAGAGCCCATCATCTCCCGAGCTGTGGTCGAGCTGAGGACTGGATTGCAACAG 661
Db 560 TGGGACGGACCCCATATCTCTCGGTACATGGGGC---CCACACACTCCATGAAGACAG 616
Qy 662 AGACCTTGTCATCTCGGCTTAATGCTCCAGCCTTTGTGCGAGCTATGCTCTGAGCCCGAG 721
Db 617 AGTACTGGCCTTTTGGCTCAACGAACCTCACTATTGTAGGCTCTGCCTATGTACTGTAGA 676
Qy 722 CTGAGTGGGGGATGAAGATGGAGCAGATGAATCTTTTCTTCAAGGAGACCTCC 781
Db 677 GTGTGGGAGCTTACCGGGGACGACGACAAAGGCTACTCTTCTTCAAGGAGCGGGGAG 736
Qy 782 GAGTGTGGGACTCTCTAGCGCATCAAGGTGCCAAGAGTGGCCGAGTGTGCGGGGG 841
Db 737 TGGAGTCCGACTGTATCCCGAGCAGGTGTGGCTGTGTCGCCGTCTGCAAGGGCG 796
Qy 842 ACCTTGGGGGAGGAAAGCCCTTACAGCAGATGGAGCAGCTTTCTTGAAGGCTGACCTGC 901
Db 797 ATATGGGGGGGCGACGGACCCCTGCAGAGGAAAGTGACACCTTCTCTGAAGCGCGGCTGG 856
Qy 902 TGTGCCAGGGCCCGAGCATGGCCGGGCTCGGGGGTCTCGAGGCTATGCGAGGCTTC 961
Db 857 CATGCTCTGCCCCGAACCTGGCAGCTCTACTTCAACAGCTGCAGGGCGATG---CACACCC 913
Qy 962 GGCCTCAGCCCTGGAGCGGGAACCCCATCTTTTATGGGATCTTTTCCCTCCAGTGGGAAG 1021
Db 914 TGCAGGACACCTCTCTGGGCACACACCTTCTTTGGGGTTTTCACGACAGTGGGGTG 973
Qy 1022 GAGCTGCCATCTCTGCTGTGTGCTGTCGACCCCAAGACATCCGGGAGTGTGTAATG 1081
Db 974 ACATGTACTGTGCGCCATCTGTGAGTACCAGTTGGAAGAGATCCAGCGGGTGTGAGG 1033
Qy 1082 GTCCCTTTAGAGACTAAACATGACTGCAACAGGGGACTGCCTGTCTATGACACACGAGG 1141
Db 1034 GCCCTATAGAGAGTACCATGAGGAAGCCCAAGAGTGGGACCGCTACACACTGACCCCT---G 1090
Qy 1142 TGCCCGACCCGAGCTGGAGAGTGCATCGCCCAACACATGAAGCTCCAGCAGTTTGGAT 1201
Db 1091 TACCCAGCCCTCGGCCCTGGCTCGTGCATTAACTGGCATCGGCCACCGCTACACCA 1150
Qy 1202 CTTACTCTCTCGTGCACACCGGTGCTCAGCTTTATCAGAGACACCCCTCTCATGAGACA 1261
Db 1151 GCTCCTGGAGTACCCGACACATCTCTCAACTTGTCTCAAGAAGCACCCGCTGTATGGAGG 1210
Qy 1262 GSCCGTGTTCGCGCTGACGGCGCCCGCTGCTGCTGCTACAGATACAGCTATCTCA 1321
Db 1211 AGCAGTGGGCTCGGTGGAGCCGCCCTGCTGCTGTAAGAAGGGCACCACTTCACCC 1270
Qy 1322 GAGTGTGGCCACAGGCTGACAGCCCTCTCAGGGAAGAATATGACGTGCTCTACCTGG 1381
Db 1271 ACCTGTGGCGACCGGTTACAGGACTTGATGGAGCCACCTATACAGTGTGTTCAATG 1330
Qy 1382 GGACAGAGATGGACACCTCCACCGGCTGTGCGCATTTGGAGCTAGCTCAGTGTCTTGG 1441
Db 1331 GCACAGGAGGGCTGGCAGCTCAAGGCTGTGAGCCTGGGGCCCTGGGTTACACCTGATTG 1390
Qy 1442 AGGATCTGCGCTGTT 1457
Db 1391 AGGAGCTGCACTGTT 1406
```

RESULT 8

```
PCT-US02-36071A-168
; Sequence 168, Application PC/TUS0236071A
; GENERAL INFORMATION:
; APPLICANT: Sagres Discovery
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000142
```

```
; CURRENT APPLICATION NUMBER: PCT/US02/36071A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: US 10/052,482
; PRIOR FILING DATE: 2001-11-08
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 168
; LENGTH: 2589
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US02-36071A-168

Query Match          9.2%; Score 215.2; DB 1; Length 2589;
Best Local Similarity 48.9%; Pred. No. 9.7e-49;
Matches 747; Conservative 0; Mismatches 748; Indels 33; Gaps 5;

Qy 136 CCCAGAACCTCGCTGCCCATCTCCGAGGCTGACTCTATCTACCCCGTTCGACGCGTCT 195
Db 76 CCCATACCCCGATCACCTGGGAGCACAGAGAGTGCACCTGGTGCAGTTCATCAGGCCA 135
Qy 196 CATAGTACAATTTACTCTGCTCTCTCTTGTGGATCTCTCCCTCCACACACTTTTACGTCGT 255
Db 136 GACATCTAACAATACTACTAGCCCTTGTCTGCTGACCGAGGACAAGGACACCTTGTACATAGT 195
Qy 256 GCAGGGATAGCATCTTCGCTTTTAACTCCCTCCCTCTCTGGGGAAGACCCCGAAGATC 315
Db 196 GCCGGGAGGGGCTTTCGCTGTGAACGACTCAACATCTCCGAGAAGCAGCATGAGGTG 255
Qy 316 GACTGGATGTTACCTGAGACTCAGACAGAACTGCAGGAAGAAAGGAAG---AAACAG 372
Db 256 TATTGGAAGGCTTCAGAAGACAAAAGCAAAATGTGCAGAAAGGGGAATCAAAACAG 315
Qy 373 GACAAATCTCAAAATTTTATCCAGATTCCTGCCAATTCGCTCTCACTCTCACTCTCACG 432
Db 316 ACAGAGTCCCTCAACTACATCCGGTGTCTGCAGCCACTCAGCGCCACTTCCCTTTACGTG 375
Qy 433 TCGGCGACTTTCGCTTTTGTATCCGAAGTGCAGGGTATTATGATGTGTCCAGTTTCCAGCAG 492
Db 376 TGTGGGACCAACGATTCAGCGCGGCTGTGACCACTGAACTTAACTCTCTTTAAGTTT 435
Qy 493 GTTGAAGACTTTGAGAGCGCGGGGAAATGTCTTTTGAGCCAGCTCAACGGTCACGA 552
Db 436 CTGGGAAATAATGAAGATGGCAAGGAAGATGTCCCTTTGACCCAGCAGCTACACA 495
Qy 553 GCTGTAATGGTGGGGCGCTCTCTACACCCCACTGTGAAGAACTTCTCGGGGACTGAG 612
Db 496 TCCGTGATGTTGATGGAGAACTTATTCGGGACGCTGCTATAATTTTGGGAAGTGA 555
Qy 613 CCCATCATCTCCGAGCTGTGGTGCAGCTGAGGACTGGATTCGAAACAGACCTTGTCA 672
Db 556 CCCATCATCTCCGAAATTTCTCCACAGTCTCTGAGGACAGAAATATGCAATCC---- 610
Qy 673 TCCTGGCTTAATGCTCCAGCCTTGTTCGAGCTATGTTCTGAGCCAGCTGAGTGGGG 732
Db 611 -CTTGGCTGAACAGCCCTAGTTTGTGTGTTGCTGACGTGATCCGAAAAGCCAGCAGC 669
Qy 733 GATGAAGATGGAGCAGTGAATCTTTTCTTTCACGAGACCTCCCGAGTGTGGAC 792
Db 670 CCCGAGCGGAGGATGACAGGGTCTACTTCTTCTTTCACGAGGTGCTGTGTGAGTATGAG 729
Qy 793 TCCTATGAGCGCATCAAGTCCCAAGAGTGGCCCGAGTGTGTGGGGGACCTTGGGGGC 852
Db 730 TTTGTGTTTCAGGGTCTGATCCCGGATAGCAAGAGTGTGCAAGGGGACCGAGCGGC 789
Qy 853 AGGAAGACCTTCAGCAGAGATGGACGAGTTCCTGAAGGCTGACCTGCTGTGCCCGAGG 912
Db 790 CTGAGGACCTTGCAGAGAAATGGACCTCTCTTGAAGAGCCGCACTCATCTGCTCCCGG 849
Qy 913 CCCGAGCATGGCCGGCCCTCCGGGTTCTGCAAGGCTATGGCAGAGCTTCGGGCTCAGCT 972
Db 850 CCAGACACGCGCTTGGTCTTCAATGTGTGCGGGATGCTCTGCTGCTCAGGTG---CCCG 906
Qy 973 GGAGCGGAACCCCATCTTTTATGGGATCTTTTCTCTCCAGTGGGAAGAGAGCTGCCATC 1032
```

Db 907 GGCTGAAGGTGCCTGTTCTATGCACTCTTCAACCCACAGCTGAACAAGCTGGGGCTG 966
Qy 1033 TCTGCTGTGTGGC-----TTCGGACCCCAAGACATCCGGGGAGTCTGAATGGT 1083
Db 967 TCGGAGTGTGGCTTACAACTGTCCACAGCCGAGGAGTCTTCTCCACGGGAAGTAC 1026
Qy 1084 CCCTTTAGAGAGCTAAACATGACTGCAACAGGGGACTGCCTGTGTCATGGACAAGAGGTG 1143
Db 1027 ATGCAGAGCACACAGTGGAGCAGTCCACACCAAGTGGGTGGGTATAAATGGCCCGTA 1086
Qy 1144 CCCAGCCAGACCTGGAGAGTGCATGCCAACACATGAAGCTCCAGCAGTTTGGATCC 1203
Db 1087 CCCAAGCCGGCCCTGGAGCGTGCATCGACAGCGAGGACCGGGCCGCAACTACACCCAGC 1146
Qy 1204 TCACCTCTCCCTGCCAGACCGGTGCTCACCTTTATCAGAGCACCCCTCTCATGACACAGG 1263
Db 1147 TCCTTGAATTTGCCACAGACAGAGCTGCAGTTCGTTAAAGACACCCCTTTCATGGATGAC 1206
Qy 1264 CCGGTGTTCCGGCTGACGGCCGCCCTGCTGCTACTACAGATACAGCCCTATCTCAGA 1323
Db 1207 TCGGTAACCCCAATAGACAACAGGCCAGGTTAATCAAGAAAGATGTGAACCTACACCCAG 1266
Qy 1324 GTCGTGGCCACAGGTCACAGCTCTCAGGGAAGAAATATGAGTGCTCTACCTGGGG 1383
Db 1267 ATCGTGTGGACCGGACCCAGGCCCTGGATGGGACTGTCTATGATGTCATGTTGTCAGC 1326
Qy 1384 ACAGAGGATGACACCTCCACCGGCTGTGCGCATTTGGAGCTCAGTGTCTTGGAG 1443
Db 1327 ACAGACGGGAGCTCTGCAAAAGCCATAGCTTCGAGCAGCTGTTCACATCATCAG 1386
Qy 1444 GATCTGGCTTTTCCCAAGAACACAGCCGGTGTAGAGCATGAATTTGA 1493
Db 1387 GAGACCCAGCTCTTCCAGGACTTTTGAGCCAGTCCAGACCCCTGCTCTCTTCAAAGAAG 1446
Qy 1494 --CCAGATTGGCTCTGGTGGCTCCCATCTACTGAGGTGACACAGTGAACACAGAAC 1551
Db 1447 GGCAACAGGTTGTCTATGCTGGCTCTAACTCGGGCGTGGTCCAGGCCCTGGCCCTTC 1506
Qy 1552 TGTGGCCTCTCCAGAGCTGCTCGGAGTGTATCTGCGCCAGGACCCGCTGTGCGCCCTGG 1611
Db 1507 TGTGGGAAGCAGGCACTGCGAGACTGTGTGTCGGCGGACCCCTACTGCGCCTGG 1566
Qy 1612 AGCTTCCGGCTGATGCTTGTGTGSCC 1639
Db 1567 AGCCCGCCACAGCAGCTGCGTGGCTC 1594

RESULT 9
PCT-US02-36071A-167
; Sequence 167, Application PC/TUS0236071A
; GENERAL INFORMATION:
; APPLICANT: Sagres Discovery
; APPLICANT: Morris, David W.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000142
; CURRENT APPLICATION NUMBER: PCT/US02/36071A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: US 10/052,482
; PRIOR FILING DATE: 2001-11-08
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 167
; LENGTH: 4157
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US02-36071A-167

Query Match 9.2%; Score 215.2; DB 1; Length 4157;
Best Local Similarity 48.9%; Pred. No. 1.2e-48;
Matches 747; Conservative 0; Mismatches 748; Indels 33; Gaps 5;

Qy 136 CCCAGAACCTCGCTGCCCATCTCCGAGGCTGACTCTCTATCTACACCCGGTTTGCAGCGTCT 195
Db 163 CCCATACCCCGATCACCTGGGAGCACAGAGAGTGCACCTGGTGCAGTTTTCATGAGCCA 222
Qy 196 CATAGTACAATTAATCTGCTCTCTCTCTGCTGATCCTGCTCCACACACTTTACGTCGGT 255
Db 223 GACATCTACAATTAATCTGCTGCTGAGCGAGACAAAGGACACCTTGTACATAGGT 282
Qy 256 GCAGGGATAGACTTTCGCTTTAACCTTCCTCTCTGGGGAAGACCCCGAAGATC 315
Db 283 GCCGGAGGCGGTCTTCGCTGTGAACCACTCAACATCTCCGAGAAGCAGCATGAGGTG 342
Qy 316 GACTGGATGTGTAACCTGAGACTACAGACAGAACTGACAGAAAGGCAAG---AAAGAG 372
Db 343 TATTGGAAGGTCTCAGAAAGACAAAAGCAAAATGTCAGAAAGGGGAAATCAAAACAG 402
Qy 373 GACGATGTCACATTTTATCCAGATTCTGCCCATTTGATGCTCTCACTCTCACTCTCCTCAG 432
Db 403 ACAGAGTGCCTCAACTACATCCGGGTGCTGAGCCACTCAGCGCCACTTCCTCTTACGTG 462
Qy 433 TCGGCACTCTCGCTTTTGTATCCGAAGTCCGGGTATTGATGTGTCAGTTTCCAGCAG 492
Db 463 TGTGGACCAACGCATTCACCGGCTGTGACCACTTGAACATCTCTTAAAGTTT 522
Qy 493 GTTGAAGACTTTGAGAGCGCGGGGAAATGTCTTTTGTAGCCAGCTCAACGGTCAGCA 552
Db 523 CTGGGAAAAATGAAGATGGCAAGGAAGATGTCTCTTTGACCCACACAGCTACACA 582
Qy 553 GCTGTAATGGCTGGGGGCGTCTCTACACCCGCTGTGAAGAACTTCTCTGGGACTGAG 612
Db 583 TCCGTCATGGTGTGAGGAACTTTATTCGGGGAGCTGCTATTAATTTTGGGAAGTGA 642
Qy 613 CCCATCATCTCCCGAGCTGTGGTGCAGCTGAGGACTGGATTTCGAACAGAGACCTTGTC 672
Db 643 CCCATCATCTCCGAAATCTTCCACAGTCTCTGTGAGACAGATATGCAATCC---- 697
Qy 673 TCTGGCTTAATGTCTCCAGCCTTTGTGCGAGCTATGGTCTGTAGCCAGCTGAGTGGGG 732
Db 698 -CTTGGCTGAACGAGCTTAGTTCTGTTGTCAGCTGTGTCGGAAGAACCCAGACAGC 756
Qy 733 GATGAAGATGAGACGATGAATCTTTTCTTTCAGGAGACCTCCCGAGCTTGGAC 792
Db 757 CCCGACGCGGAGGATGACAGGCTCTACTTCTTTCACGGAGTGTCTGTGGAGTATGAG 816
Qy 793 TCCATATGAGCGCATCAAGGTCCCAAGAGTGGCCCGAGTGTGTGCGGGGAGCTTGGGGC 852
Db 817 TTTGTGTTCCAGGCTCTGATCCCAGGATAGCAAGAGTGTCAAGGGGAGCCAGGCGGC 876
Qy 853 AGGAAGACCTTTCAGCAGAGATGGACAGCTTTTGAAGGCTGACCTGCTGTGCCCCAGG 912
Db 877 CTGAGGACCTTGCAGAAGAAATGGACCTCTCTCTGAAAGCCGAGCTCATCTGCTCCGG 936
Qy 913 CCCGAGCATGCCGGGCTCCGGGTTCTGAGGCTATGGCAGACTTTCGGCCTCAGCT 972
Db 937 CCAGACAGCGGCTTGGTCTCAATGTGCTGGGAGTGTCTTCTGCTCAGGTC---CCG 993
Qy 973 GGAGCGGAACCCCATCTTTTATGGGATCTTTTCTCTCCAGTGGGAAGAGCTGCCATC 1032
Db 994 GGCCTGAAGTGCCTGTGTCTATGCACTTTCACCCACAGCTGAACACAGTGGGGTG 1053
Qy 1033 TCTGCTGTGTGGC-----TTCCGACCCCAAGACATCCGGGAGTGTGTAATGTT 1083
Db 1054 TCGGAGTGTGCGCTTACAACCTGCCACAGCCGAGAGGTCTTCTCCACCGGAAGTAC 1113
Qy 1084 CCTTTAGAGCTAAACATGACTGCAACAGGGGACTGCCTGTGTCATGGACACAGGTG 1143
Db 1114 ATGACAGACCAACAGTGGAGCAGTCCACACCAAGTGGGTGCGCTATAATGCGCCGTA 1173
Qy 1144 CCCAGCCAGACCTGGAGTGCATCGCCAACACATGAAGCTCCAGCAGTTTGTATCC 1203
Db 1174 CCCAGCCGCGCTGGAGCTGTCATCGACAGGAGCCGCGGCCCACTACACCAGC 1233
Qy 1204 TCACTCTCCCTGCCAGACCGGTGCTCACCTTTATCAGAGACCAACCCCTCTCATGACAGG 1263

Db 1234 TCCTTGAATTTGCCACACAGAGCGTGCAGTTCGTTAAAGACACACCTTTTGATGATGAC 1293
QY 1264 CCCTGTGTTCCGGCTCAGCGCCGCCCTGCTGCTCAGTACAGATACAGCTTATCTCAGA 1323
Db 1294 TCCTGTAACCCCAATACACACAGCGCCAGGTAACTCAAGAAAGATGTGAATACACCCAG 1353
QY 1324 GTGCTGGCCACAGGTGACAGCTCTCAGGGAAGAAATATAGCTGTCTACCTGGGG 1383
Db 1354 ATGCTGCTGGACCGGACCCAGGCGCTGGATGGAGCTGTCTATGATGCTATGTTGTCAGC 1413
QY 1384 ACAGAGATGGACACCTCCACCGGCTGCGCATTTGGAGCTCAGTCTAGTCTTGGAG 1443
Db 1414 ACAGACGGGGAGCTGTGCACAAAGCCATCAGCGCTCGAGCAGCTTTTACATCATCGAG 1473
QY 1444 GATCTGGCTTGTTCACAGAACACAGCGGTTGAGAGCATGAATTTGTA----- 1493
Db 1474 GAGACCCAGCTCTCCAGGACITTTGAGCCAGTCCAGACCCCTGCTGCTTCAAGAAG 1533
QY 1494 --CCAGATTTGGCTCTGGTGGCTCCCATATCTAGGTGACAAAGTGACACACAGCAAC 1551
Db 1534 GGCACAGGTTTTGCTATGCTGCTTAACCTGGGCGTGTCTCAGGCCCGCTGGGCTTC 1593
QY 1552 TGTGGCCGCTCCAGAGCTGCTCGGAGTATATCTTGGCCAGGACCCCGTGGCGCTGG 1611
Db 1594 TGTGGGAAGCAGGCGACCTGCGAGGAGTGTGCTGGCGGGAGCCCTACTGCGGCTGG 1653
QY 1612 AGCTTCCGCTTGTATGCTTGTGGGCC 1639
Db 1654 AGCCGCCACAGCGACCTGCGTGGCTC 1681

RESULT 10

PCT-US02-36071A-165

; Sequence 165, Application PC/TUS0236071A

; GENERAL INFORMATION:

; APPLICANT: Sagres Discovery

; APPLICANT: Morris, David W.

; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER

; FILE REFERENCE: 52945200142

; CURRENT APPLICATION NUMBER: PCT/US02/36071A

; CURRENT FILING DATE: 2002-11-08

; PRIOR APPLICATION NUMBER: US 10/052,482

; PRIOR FILING DATE: 2001-11-08

; NUMBER OF SEQ ID NOS: 241

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 165

; LENGTH: 2586

; TYPE: DNA

; ORGANISM: Mus musculus

PCT-US02-36071A-165

Query Match 8.8%; Score 205.6; DB 1; Length 2586;

Best Local Similarity 49.2%; Pred. No. 4,3e-46;

Matches 733; Conservative 0; Mismatches 724; Indels 33; Gaps 6;

QY 174 TCTACCCCGTTTGCAGCTCTCATACGTACAAATTAATCTGCTCTCTTGTGGATCCTGC 233
Db 114 TCTGTGTCAGTTTCAAGCCAGGCATCTTTAACTACTCGGCTTGTGATGAGTGAGGA 173
QY 234 CTCACACACATTTAGTGGGTGCACGGGATAGCATCTTTCGCTTTAAACCTCCCTTCTC 293
Db 174 CAAGACACTCTGTATGTAGGCGCCGGGAAGCAGTCTTTGCAGTGAATGCGCTGAACAT 233
QY 294 TGGGAAGACCCCGAAGGATCGACTGGATGTTACTTGAGCTCAGACACAGAACTGGAG 353
Db 234 CTCCTGAGAGCAACATGAGGTATATTGGAAGCTCTGGAAGCAAAAATCCAAAGTGTGC 293
QY 354 GAAGAAGGCAAA--GAAGAAGGAGGAATGTACAAATTTTATCCAGATTTCTCGCCATGT 410
Db 294 AGAGAAGGGGAATCAAGACAGACCGGAATGCCTAACTACATTCGAGTACTACACCCACT 353

QY 411 CAATGCTCTCACCTCCTCAGCTGGCGCACCTTTCGCTTTTGTATCCGAAGTGGGGTTAT 470
Db 354 AAGCAGCACTTCCCTCTATGTGTGGACCAATGCTTCCAGCCACCACTGTGACCACT 413
QY 471 TGNATGTGTCAGTTTCCAGCAGGTTGAAGACTTTAGAGCGGCGGGGAAATGTCTCTTT 530
Db 414 GAATTTGACATCCTTCAAGTTTCTGGGAAAAGTGAAGATGGCAAGGAAGATGCCCCCT 473
QY 531 TGAGCCAGCTCAACGCTCAGCAGCTGTAATGGCTGGGGCGTCTCTACACCCCACTGT 590
Db 474 CGACCCCGCCACAGCTACACATCAGTCATGGTTGGGGGAGCTCTACTCTGGGAGCTC 533
QY 591 GAAGAATTTCTGGGAGCTAGCCCCATATCTCCAGCTGTGGGTGCGAGCTGAGGAGTG 650
Db 534 CTATATTTCTTGGGAGTGAACCCATCATCTCTCGAAACTCTTCCACAGTCCCTTGAG 593
QY 651 GATTCGAACAGAGACCTTGTCTATCTCTGGCTTAATGCTCCAGCTTTTGTCCAGCTATGT 710
Db 594 G-----ACGGAGTATGCCATCCGTTGGCTGAACAGAGCTAGTCTCTTGTCTGACGT 647
QY 711 CCTGAGCCCACTGAGTGGGGGATGAAGATGGAGACGATGAATCTTTTCTTCTTCTCAC 770
Db 648 GATCCAGAAAAGCCCAAGATGGTCCGAGGGTGAAGATGACAAGTCTACTTCTTTTAC 707
QY 771 GGAGACCTCCCGAGTGTGGACTCTTATGAGCGCATCAAGGTCCCAAGAGTGGCCGAGT 830
Db 708 GGAGGTATCCGTGGAGTACGAATTCGTCTTCAAGTTGATGATCCCGGAGTGGCCAGGT 767
QY 831 GTGTGGGGGAGCTTGGGGGAGGAGACCTTTCAGCAGAGATGGACACGTTTCTTGAA 890
Db 768 GTGCAAGGGCGACAGGCGGCTGCGGACTTTGCAAAAAGATGGACCTCTCTCTTAA 827
QY 891 GGCTGACCTGCTGTGCCAGGGCGGAGCATGCGCGGCTCCGGGTTCTTGACGCTAT 950
Db 828 GCGCAGGCTGATCTGCTCAAGCCACAGAGTGGCTGGTCTTCAACATACTTCAGATGT 887
QY 951 GGCAGAGCTTCGGCTTCAGCTGGAGCGGAACCCCATCTTTATGGAGTCTTTTCTC 1010
Db 888 GTTGTGCTGAGGGC---CCCGGCTCAAGAGAGCTGTGTTCTATGCGGTCTTCAACC 944
QY 1011 CCAGTGGGAAGAGCTGCCATCTGCTGTGTGCTGCTTCCGACCCCAAGACATCCGGC 1070
Db 945 ACAGCTGAACAATGTGGGTCTGTGAGCGGTGCGCTACACACTGGCCACGCTGGAGGC 1004
QY 1071 AGT-----GCTGAATGCTCTTTAGAGAGCTAAA--ACATGACTGCAACAGGGGACT 1121
Db 1005 AGTCTTCTCCCGTGGAAAGTACATCAGAGTGCCACAGTGGAGCAGTCTCACACCAAGTG 1064
QY 1122 GCCTGTGACAGACAGAGGTGCCCGAGCCAGACCTGGAGAGTGGATCGCCAAACAT 1181
Db 1065 GGTGGGCTACAATGGCCAGTGGCCACTCCCGAGCTGGAGGCTGTATCGACAGTGGGC 1124
QY 1182 GAAGCTCCAGAGTTTGGATCTCACTCTCCCTGCCAGACCGGCTGCTCACCTTTATCAG 1241
Db 1125 CCGGGAGCCAACTACACAGCTCTTGAATCTCCAGACAAACACTGCAGTTGTGAAA 1184
QY 1242 AGACACCCCTCTCATGGAGCGGCTGTTCGCGGTGAGCGCGCCCTCTGTGCTAC 1301
Db 1185 AGACACCCCTTTGATGGATGACTCAGTGACCCCGATAGACACAGACCCCAAGCTGATCAA 1244
QY 1302 TACAGATACAGCTATCTCAGAGTGTGGGCCACAGGGTGGACCGCTCTCAGGGAAGA 1361
Db 1245 AAAAGATGTAACTACACCCAGATAGTGGTAGACAGGACCCAGGCGCTGGATGGGACTTT 1304
QY 1362 ATATCAGTGTCTACCTGGGACAGAGATGGACACCTCCACCGGCTGTCGCAATGG 1421
Db 1305 CTACAGCTGATGTTTATCAGCAGACAGCCGGGAGCTGTGCATAAGAGTCTATCTTAC 1364
QY 1422 AGCTCAGCTCAGTGTGCTTTGGAGATCTGGCCCTGTTTCCAGAACACACAGCCGTTGAGAG 1481
Db 1365 AAAAGAGGTGCTATCGAGGAGACCCCAACTCTTCGCGGACTTTGAACCGGCTCTAAC 1424
QY 1482 CATG-----AAATGTACCAGATTGGCTCTCTGGTGGGCTCCCACTACTGAGGT 1529

Db 1425 TCTGCTGCTATCGTCAAGAAGGGGAGGAAAGTTTGTCTATGACGAGCTCCAACTCTGGAGT 1484
QY 1530 GACACAACTGAACACACCACTGGCGCTCCAGAGCTGCTCGGAGTGTATCCTGGC 1589
Db 1485 GGTCAAGCGCCCTGGGATCTCTCGAAGACACGGTAGCTGTGAAGACTGTGTGTAGC 1544
QY 1590 CCAGACCCCGTGTGGCGCTGAGCTTCCGGCTTGATGCTTGTGTGGGCC 1639
Db 1545 ACGGAGCCCTACTGTGCTGGAGCCAGCCATCAAGGCCCTGTGTACCC 1594

RESULT 11

PCT-US02-36071A-164
; Sequence 164, Application PC/TUS0236071A
; GENERAL INFORMATION:
; APPLICANT: Sagres Discovery
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000142
; CURRENT APPLICATION NUMBER: PCT/US02/36071A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: US 10/052,482
; PRIOR FILING DATE: 2001-11-08
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 164
; LENGTH: 2769
; TYPE: DNA
; ORGANISM: Mus musculus
PCT-US02-36071A-164

Query Match 8.8%; Score 205.6; DB 1; Length 2769;
Best Local Similarity 49.2%; Pred. No. 4.4e-46;
Matches 733; Conservative 0; Mismatches 724; Indels 33; Gaps 6;

QY 174 TCTACCCGGTTTGACGGGTCTCATACGTACAAATTTACTCTGCTCTCTCTGTTGGATCCCTGC 233
Db 174 TCTGTCGAGTTTCAAGCCAGGCGATCTTTAACTACTCGGCCCTTGCTGATGAGTGAGGA 233
QY 234 CTCCACACACTTTACGTCGGTGCACGGGATGATCTCTGCTTTTAAACCTCCCTCTCTC 293
Db 234 CAAGACACACTCTGTATGTAGGCGCCGGGAGCAGCTCTTGCAAGTGAATGCGCTGAACAT 293
QY 294 TGGGAAAGACCCGAGGATCGACTGTGATGATGATGATGATGATGATGATGATGATGATG 353
Db 294 CTCTGAGAGCAACATGAGGTATATTGGAAGTCTCTGAAGACAAATAATCCAAAGTGTGC 353
QY 354 GAAGAAAGGCAA--GAAGAGGAGCAATGTACACAATTTTATCCAGATTTCTGCCATTTGT 410
Db 354 AGAGAAGGGGAAATCAAGACAGACGGAATGCTTAACTACATTTCTGAGTACTACACCCACT 413
QY 411 CAATGCCCTCCTACCTCTCAGTGGGACCTTCCGCTTTGATCGAAGTGGGGGTTAT 470
Db 414 AAGACACACTTCCCTCTATGTGTGGGACCAATCGTTCCAGCCACCTGTGACCACT 473
QY 471 TGATGTGCCAGTTTCCAGCAGGTTGAAGACTTCAGAGCGCGCGGGGAATGTCCTT 530
Db 474 GAACTTGACATCTTCAAGTTTCTGGGAAAGTGAAGATGAAGATGAAGATGAAGATGCCCCTT 533
QY 531 TGAGCCAGCTCAACGGTCAGACGCTGTAAATGGCTGGGGCGTCTCTTACACCGCCACTGT 590
Db 534 CGACCCGCCACACACTACACATCAGTCATGTTGGGGCGAGCTCTACTCTGGGACGTC 593
QY 591 GAAGAACTTCTCTGGGAGTGAAGCCCATCATCTCCGAGCTGTGGGTGAGCTGAGACTG 650
Db 594 CTATAATTTCTGGGAGTGAAGCCCATCATCTCTGAAACTCTTCCACAGTCCCTTGAG 653
QY 651 GATTCGACAGACAGACTTGTCTATCTCTGCTTAATGCTCCAGCCTTTGTGCGAGTATGGT 710
Db 654 G-----ACGGAGTATGCCATCCCGTGGCTGAACGAGGCTAGCTTCGCTCTTGTGCTGAGCT 707

QY 711 CCTGAGCCAGCTGAGTGGGGGATGAAGATGAGACGATGAGACGATGAGAACTTTTTTTTCTTAC 770
Db 708 GATCCAGAAAAGCCCGAGATGGTCCGAGGGTGAAGATGACAAGTCTACTTCTTTTAT 767
QY 771 GAGACCTCCCGAGTGTGGACTCTATGAGGCGATCAAGGTCCCAAGAGTGGCCGAGT 830
Db 768 GSAGGTATCCGTGGAGTACGAATTCGTCTCAAGTTGATGATCCGCGAGTGTCCAGGCT 827
QY 831 GTGTCCGGGGACCTTGGGGGAGGAAGACCTTCCAGCAGAGATGAGACGATGTTCTGAA 890
Db 828 GTGCAAGGGCGACAGGGGGCTCGGACTTTGCAAAAAGTGGACCTCTTCCCTAAA 887
QY 891 GGCTGACCTGTGTGCCAGGGCCGAGCATGGCGGGCTCCGGGTTTCTCAGAGCTAT 950
Db 888 GGCCAGGCTGATCTCTCAAGCCAGACAGTGGCTGGTCTTCAACATACTTTCAGAGTCT 947
QY 951 GGCAGAGCTTCCGGCTCAGCTGGAGGGGAAACCCCATCTTTTATGGATCTTTTCTCTC 1010
Db 948 GTTGTGTGTAGGGC---CCCGGGCTCAAGGAGCTGTGTCTATGCGGTCTTCAACCC 1004
QY 1011 CCAGTGGGAAGGAGCTGCCATCTCTGCTGTGTGTGCTTCCGACCCCAAGACATCGGGC 1070
Db 1005 ACAGCTGAACATGTGGTCTGTACGGGTGTGCGCTACACACTGGCCACGGTGAGGC 1064
QY 1071 AGT-----GCTGAATGTCCTTTAGAGAGCTAAA--ACATGACTGCAACAGGGGACT 1121
Db 1065 AGTCTTCTCCCGTGGAAAGTACATGCAGAGTCCACAGTGGAGAGCTCTCACACCAAGTG 1124
QY 1122 GCCTGTATGGACAACAGAGTGGCCCGCCAGCTGGAGAGTGGAGTGCATGCCCAACACAT 1181
Db 1125 GGTGCGCTACATGGCCAGTCCCACTCCCGAGCTGGAGAGTGTATGCAGAGTGAGGC 1184
QY 1182 GAACTCTCAGCAGTTTGGATCCTCACTCTCCTGCGCAGACCCGCTGCTCACTTATCAG 1241
Db 1185 CGGGCAGCAACTACACAGCTCTTGAATCTCCAGACAAAACACTGCAGCTTTGTAAA 1244
QY 1242 AGACACACCTCTCATGGACAGCGCGTGTTCGCGCTGACGGCGCGCCCTCTGCTGTAC 1301
Db 1245 AGACACACCTTTGATGGATGACTCAGTGACCCCGATAGACAACAGACCCCAAGCTGATCAA 1304
QY 1302 TACATATACAGCTTATCTCAGAGTCTGTGGCCCAAGGTGACAGCTCTCAGGGAAAGA 1361
Db 1305 AAAAGATGTAACTACACCAAGATAGTGTAGACAGGCCAGCCCTGGATGGGACTTT 1364
QY 1362 ATATCAGCTGTCTTACCTGGGACAGAGTGGACACCTCCACCGGCTGTGCGATTCG 1421
Db 1365 CTACAGCTCATGTTTATCAGACAGACCGGGAGCTCTGCATATAAGCAGTCTATCTTAC 1424
QY 1422 AGCTCAGCTCAGTGTCTTTGGAGATCTGGCCTTTGTTCAGAACCAACAGCCGTTGAGAG 1481
Db 1425 AAAAGAGTGTGATCTCATCGAGGAGACCCCACTCTTCCGGGACTTTGAAACCGTCTTAAC 1484
QY 1482 CATG-----AAATGTACACGATTTGGCTCTCTGTTGGGTGCGCTCCCATCTAGCT 1529
Db 1485 TCTGTGTCTATCTCAAGAGAGGGAGGAAGTTTGTCTATGAGGCTCCCAACTCTGGAGT 1544
QY 1530 GACACAAAGTGAACACAGCAACTGTGGCGCTCTCCAGAGCTGCTCGAGTGTATCTCTGGC 1589
Db 1545 GTTCCAGGCCCTCTGGCATTTCTGGAAAGACAGGTAAGTGTGAAGACTGTGTGTAGC 1604
QY 1590 CCAGGACCCGCTGTGCGCTGGAGCTTCGGGTTGATGCTTGTGTGGGCC 1639
Db 1605 ACGGAGCCCTACTGTGCTGGAGCCAGCCATCAAGGCTCTGTGTATCCC 1654

RESULT 12
PCT-US03-27411-44
; Sequence 44, Application PC/TUS0327411
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; Lee, Soo Yeun;
; APPLICANT: ELLIOTT, Vicki S.; HAFALIA, April J.A.;
; APPLICANT: BURFORD, Neil; SPRAGUE, William W.;
; APPLICANT: GRIFFIN, Jennifer A.; YANG, Yonghong G.;

APPLICANT: BECHA, Shanya D.; KHARE, Reena;
APPLICANT: THORNTON, Michael B.; MASON, Patricia M.;
APPLICANT: GIETZEN, Kimberly J.; ISON, Craig H.;
APPLICANT: MARQUIS, Joseph P.; SWARNAKAR, Anita;
APPLICANT: RAMKUMAR, Jayalaxmi; JIN, Pei;
APPLICANT: RICHARDSON, Thomas W.; TRAN, Oyen K.
TITLE OF INVENTION: NEUROTRANSMISSION-ASSOCIATED PROTEINS
FILE REFERENCE: PF-1557 PCT
CURRENT APPLICATION NUMBER: PCT/US03/27411
CURRENT FILING DATE: 2003-09-02
PRIOR APPLICATION NUMBER: US 60/408,383
PRIOR FILING DATE: 2002-09-04
PRIOR APPLICATION NUMBER: US 60/408,781
PRIOR FILING DATE: 2002-09-05
PRIOR APPLICATION NUMBER: US 60/414,221
PRIOR FILING DATE: 2002-09-26
PRIOR APPLICATION NUMBER: US 60/426,483
PRIOR FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 60/431,566
PRIOR FILING DATE: 2002-12-06
PRIOR APPLICATION NUMBER: US 60/434,317
PRIOR FILING DATE: 2002-12-17
PRIOR APPLICATION NUMBER: US 60/437,763
PRIOR FILING DATE: 2003-01-03
NUMBER OF SEQ ID NOS: 60
SOFTWARE: PERL Program
SEQ ID NO 40
LENGTH: 725
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: 7520844CB1
PCT-US03-27411-40

Query Match 6.2%; Score 145; DB 1; Length 725;
Best Local Similarity 73.1%; Pred. No. 1.2e-29;
Matches 302; Conservative 0; Mismatches 60; Indels 51; Gaps 7;
QY 151 CCCATCTCCGAGGCTGACTCTCTACCCGGTTTGACGGTCTCTACATACGTACAAATTAC 210
DB 273 CTCTGCTGAGAGGCTGACTCTCTGCTCACCCGGTTTCGAGTCCCTCACACATACAAATTAC 332
QY 211 TCTGCTCTCTTGGGATCTGCTCCACACACTTACGTGGGTGACGGGATAGCATC 270
DB 333 TCTGTTCTCTTGGGATCTGCTCCACACACTTATGTTGGGCGCCGGACACCATC 392
QY 271 TTGCTTTTA-ACCCTCCCTTCTCTGGGGAAGACCCCGAAGG----- 312
DB 393 TTGCTTTTATCCCTGCCCCCTTCTCAGGGGAGAGACCCCGCAGGGTGAGAGACAAGAGAGG 452
QY 313 -----ATCGACTGGATGTTACTGAGACTCAGACACAGAA 347
DB 453 GAACGGACCCCTGACCCTGTAGCATATTGACTGATGTTCTTGAGGCTCAGACACAGAA 512
QY 348 CTGAGGAAAG- AAGGCAAGAAAGAGGACGAATG-TCACAAATTTATCCAGATTC--TCG 403
DB 513 CTGTAGGAGAAAAGGCGACGACAGACGAATGTTCAAAATTTGTCCAGATTCCTCGC 572
QY 404 CCATGTCAATGCCTCTCACTCCT- ---CAGTGGGACACCTTGCCTTTTATCGGAAGT 460
DB 573 CATTTGGCAATGCTCTCACTCTTCACTTTGTGGCAACCTTTGCTTTTATCGGAAGT 632
QY 461 GCGGGTTTATTTGATGTTG-CCAGTTTCCAGCAGGTTGAAAGCTTGAGAGCGG 512
DB 633 GCGGGTTTATTTGATGTTGTCACAGGTTCCAGAGGGTTCAACAGACTTGAGGAG 685

RESULT 14
PCT-US03-09929-21
Sequence 21, Application PC/TUS0309929
GENERAL INFORMATION:
APPLICANT: Curagen Corporation, et al.

TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND MET
FILE REFERENCE: 21402-573B-061
CURRENT APPLICATION NUMBER: PCT/US03/09929
CURRENT FILING DATE: 2003-04-01
PRIOR APPLICATION NUMBER: 60/368,996
PRIOR FILING DATE: 2002-04-01
PRIOR APPLICATION NUMBER: 60/369,980
PRIOR FILING DATE: 2002-04-04
PRIOR APPLICATION NUMBER: 60/370,381
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: 60/370,969
PRIOR FILING DATE: 2002-04-08
PRIOR APPLICATION NUMBER: 60/371,002
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 60/372,002
PRIOR FILING DATE: 2002-04-12
PRIOR APPLICATION NUMBER: 60/384,297
PRIOR FILING DATE: 2002-05-30
PRIOR APPLICATION NUMBER: 60/386,816
PRIOR FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: 60/389,123
PRIOR FILING DATE: 2002-06-13
PRIOR APPLICATION NUMBER: 60/402,207
PRIOR FILING DATE: 2002-08-09
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 179
SOFTWARE: CuraSeqList version 0.1
SEQ ID NO 21
LENGTH: 1327
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (2)..(1327)
PCT-US03-09929-21

Query Match 2.7%; Score 62; DB 1; Length 1327;
Best Local Similarity 47.6%; Pred. No. 1.2e-06;
Matches 280; Conservative 0; Mismatches 275; Indels 33; Gaps 2;
QY 332 AGACTCAGACAGAACTGCAGGAAGGAAGCAAGAAAGGACCAATGTCACAATTTTA 391
DB 147 AGCCGATGTAGACACATGCAAGATGAAGGAAACATAAGATGAGTGCCACAACTTTA 206
QY 392 TCCGATTTCTGCCATTGTCAATGCTCTCACCTCTCTACGTCGCGCACCTTCGCTTTTG 451
DB 207 TTAAGTTCTCTTAAGAAACAGATGATGATTTGTTCTGTGGAACATATGCTTCA 266
QY 452 ATCCGAAAGTGGGGTTATTGATGTGCCAGTTTCCAGAGGTTGAAAGACTTTGAGAGCG 511
DB 267 ACCCTTCTCGAGAAACTATAAGATGGATACATTGGAACCATTCGGGGATGAATTCAGCG 326
QY 512 GCGGGGGAATGTCCTTTTGAGCCAGCTCAACGTCAGCAGCTGTAATGGCTGGGGGG 571
DB 327 GAATGGCCAGATGCCCATATGATGCCAACATGCCAAGCTTGCACATGTTGCGAGATGGA 386
QY 572 TCCTCTACCGCCACTGTGAAGAACTTCCTGGGGACTGAGCCCATCATCTCCCGAGCTG 631
DB 387 AACTATACTCAGCCACAGTACTGACTTCTTGGCCATTTGACCCAGTCAATTTACCGAGTC 446
QY 632 TGGGTCGAGCTGAGGACTGGATTTCGAACAGAGACCTTGTCACTCCCTGGCTTAATGCTCCAG 691
DB 447 TTGGAGAAAGCCCTACCCCTGCGGACCGTCAAGCAGCATTTCAAATGTTGAAAGAACCAT 506
QY 692 CCTTTGTCGAGCTATGTTCTTGACCCAGCTGAGTGGGGGATGAAGATGGAGAGCATG 751
DB 507 ACTTTGTTCAAGCCGTG-----GATTACGGAGATT 536
QY 752 AAATCTTTTTTTTCTTACGGAGACCTCCCGAGTGTGGAGTCTTGGACTCTCTATGAGCGCATCAAGG 811
DB 537 ATATCTACTTCTTCTTACGGGAAATAGCAGTGGAGATATAACACCATGGGAAGAGGTAGTTT 596
QY 812 TCCCAAGAGTGGCCCGAGTGTGTGCGGGGAGACCTTTGGGGGAGGAGAACCCCTTTCAGCAGA 871

```
Db      597  TCCCAAGAGTGCTCAGGTTTGAAGATGATATGGAGGATCTCAAGAGTCTGGAGA 656
QY      872  GA---TGGACGACGTTTCTGAAGGCTGACCTGCTGTGCCCGAGGCCG 916
Db      657  AACAGTGGACGTCGTTCTCTGAAGGCGCGCTTGAACGTGCTCAGTTCTG 704
```

RESULT 15

```
PCT-US03-17412-21
; Sequence 21, Application PC/TUS0317412
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation
; APPLICANT: Alvarez, Enrique
; APPLICANT: Anderson, David W.
; APPLICANT: Dhanabal, Mohanraj
; APPLICANT: Khrantsov, Nikolai V.
; APPLICANT: Larochelle, William J.
; APPLICANT: Li, Li
; APPLICANT: Lichenstein, Henri
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: SEMAPHORIN-LIKE PROTEINS AND METHODS OF USING SAME
; FILE REFERENCE: 15966-540CIP2
; CURRENT APPLICATION NUMBER: PCT/US03/17412
; CURRENT FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: 09/520,781
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/123,667
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: 60/234,082
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 60/233,798
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 60/174,485
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: 10/403,676
; PRIOR FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/384,798
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/402,407
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/443,062
; PRIOR FILING DATE: 2003-01-28
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 21
; LENGTH: 1327
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1327)
PCT-US03-17412-21
```

```
Query Match      2.7%; Score 62; DB 1; Length 1327;
Best Local Similarity 47.6%; Pred. No. 1.2e-06;
Matches 280; Conservative 0; Mismatches 275; Indels 33; Gaps 2;

QY      332  AGACTCACAGACAGAACTGCAGGAAGAAAGGACGACGAATGTCAACAATTTTA 391
Db      147  AGGCCGATGTAGACACATGCAGAAATGAAGGAAACATAAGGATGAGTGCCCAACTT 206
QY      392  TCCAGATTCTGCCCAATGTCAATGCCTCTCACGTGCGGACCCCTTCGCTTTTG 451
Db      207  TTAAGATTCTTCTAAAGAAAACGATGATGCAATTTGTCGTGGAACATAATGCCTCA 266
QY      452  ATCCCAAGTGGCGGGTTATTGATGTCTCCAGTTCACAGAGGTTGAAAGACTTGAGAGCG 511
```

```
Db      267  ACCCTTCTTCGAGAAACTATAAGATGGATACATTGGAACCAATTCGGGGATGAATTCAGCG 326
QY      512  GCCGGGGAAATGTCTCTTTTGAAGCCAGCTCAACGGTCACAGCTGTAAATGGCTGGGGCG 571
Db      327  GAATGCCAGATGCCCATATGATGCCAAACATGCCAACGTTGCACCTGTTTGCAGATGGAA 386
QY      572  TCCTCTACACCGCCACTGTGAAGAACTCTCTGGGACTGAGCCCATCATCTCCCGAGCTG 631
Db      387  AACTATACTCAGCCACAGTGACTTCTCTTGGCCATTGACGACGTCATTTACCGGAGTC 446
QY      632  TGGTGGAGCTGAGGACTGGATTTCGAACAGAGACCTTGTTCATCTCTCCCTGCTTAATGCTCCAG 691
Db      447  TTGGAGAAAGCCCTTACCCTCGGACCGTCAAGCAGGATTCAAAATGGTTGAAAGAACCAT 506
QY      692  CCTTTGTGCGCAGCTATGTCCTGAGCCAGCTGAGTGGGGGATGAAGATGGAGACGATG 751
Db      507  ACTTTGTTCAAGCCGTG-----GATTACCGGACATT 536
QY      752  AAATCTTTTTTTTCTACGGAGACCTCCCGAGTGTGGACTCTCTATGAGCGCATCAAGG 811
Db      537  ATATCTACTTCTTCTCAGGAAATAGCAGTGGAGTATAACACCATGGGAAAGTAGTTT 596
QY      812  TCCCAAGAGTGGCCCGAGTGTGCGGGGACCTTGGGGGAGGAGAGACCCCTTCACGACA 871
Db      597  TCCCAAGAGTGGCTCAGGTTTGTAAAGATGATATGGGAGGATCTCAAGAGAGTCTCGAGA 656
QY      872  GA---TGGACGACGTTTCTGAAGGCTGACCTGTGCCCCAGGCCG 916
Db      657  AACAGTGGACGTCGTTCTCTGAAGGCGCGCTTGAACGTGCTCAGTTCTG 704
```

Search completed: September 26, 2003, 11:35:56
Job time : 365.737 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2003, 19:18:34 ; Search time 4613.83 Seconds
(without alignments)
12279.116 Million cell updates/sec

Title: US-09-284-180a-2

Perfect score: 2331

Sequence: 1 atgttgccaggccgagcg.....gtgatgagacccatctaa 2331

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:**
2: em_esthum:**
3: em_estin:**
4: em_estmu:**
5: em_estov:**
6: em_estpl:**
7: em_estro:**
8: em_htc:**
9: gb_est1:**
10: gb_est2:**
11: gb_htc:**
12: gb_est3:**
13: gb_est4:**
14: gb_est5:**
15: em_estfun:**
16: em_estom:**
17: em_gss_hum:**
18: em_gss_inv:**
19: em_gss_pln:**
20: em_gss_vrt:**
21: em_gss_fun:**
22: em_gss_mam:**
23: em_gss_mus:**
24: em_gss_pro:**
25: em_gss_rod:**
26: em_gss_phg:**
27: em_gss_vrl:**
28: gb_gss1:**
29: gb_gss2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1515	65.0	2719	11 AK035993	AK035993 Mus muscu
2	1511.8	64.9	2855	11 AK047067	AK047067 Mus muscu
3	645	27.7	704	14 CB527136	CB527136 UI-M-FY0-
4	632.6	27.1	691	12 BM944530	BM944530 UI-M-EH0p

5	608	26.1	868	10 BG333723	BG333723
6	596.4	25.6	693	12 BM963828	BM963828
7	586.2	25.1	901	13 BQ213659	BQ213659 AGENCOURT
8	560	24.0	1154	12 BM462308	BM462308 AGENCOURT
9	557.6	23.9	733	10 AW953466	AW953466 EST365536
10	547.8	23.5	765	12 BI823626	BI823626 603038515
11	546.8	23.2	733	10 AW957134	AW957134 EST369204
12	540	23.5	701	14 CB245456	CB245456 UI-M-FY0-
13	508.8	21.8	772	12 BI915256	BI915256 603184795
14	496	21.3	758	14 CB244307	CB244307 UI-M-FY0-
15	484	20.8	669	10 BB628682	BB628682 BB628682
16	475.6	20.4	578	12 BI344452	BI344452 372924 MA
17	464.2	19.9	669	12 BG774505	BG774505 602662571
18	455.6	19.5	486	10 BE982860	BE982860 UI-M-CG0p
19	437.2	18.8	466	9 AW123399	AW123399 UI-M-BH2
20	405.2	17.4	434	9 AW123577	AW123577 UI-M-BH2
21	397.4	17.0	556	10 BE683223	BE683223 182687 MA
22	393	16.9	444	13 BY208889	BY208889 BY290889
23	386.8	16.6	503	9 AA459837	AA459837 2x50g12.r
24	381.2	16.4	1045	10 BE734978	BE734978 601570573
25	380	16.3	380	14 CB810579	CB810579 AMGNNUC:Y
26	375.2	16.1	522	10 BE667719	BE667719 155505 MA
27	366.4	15.7	368	10 BF415905	BF415905 UI-R-CA0-
28	363.8	15.6	466	12 BI275275	BI275275 UI-R-CX0-
29	352.2	15.1	416	13 BY224521	BY224521
30	346.6	14.9	396	13 BY203844	BY203844 BY203844
31	340	14.6	616	14 BY731493	BY731493 BY731493
32	333	14.3	769	13 BQ745165	BQ745165 UI-M-EH0p
33	314.2	13.5	361	13 BY192163	BY192163 BY192163
34	312	13.4	430	13 BY225500	BY225500 BY225500
35	308.8	13.2	354	13 BY176252	BY176252 BY176252
36	304.2	13.1	426	10 BE754071	BE754071 207388 MA
37	303.4	13.0	365	14 CD355327	CD355327 UI-M-FY0-
38	296.2	12.7	343	13 BY186478	BY186478 BY186478
39	295.8	12.7	456	10 BE477365	BE477365 161006 BA
40	293.8	12.6	308	10 BE115296	BE115296 UI-R-CA0-
41	293.6	12.6	517	14 R54387	R54387 YG78f12.r1
42	289.6	12.4	337	14 245329	245329 HSC2MG091 n
43	287.4	12.3	364	14 T09073	T09073 EST06966 n
44	284.4	12.2	376	9 AW436751	AW436751 77209 MAR
45	283.8	12.2	437	14 H24181	H24181 ym55f02.r1

ALIGNMENTS

RESULT 1
AK035993
LOCUS
DEFINITION
AK035993 Mus musculus 16 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:9630025H20 product:sema domain, immunoglobulin domain (lg), TM domain, and short cytoplasmic domain, full insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AK035993
AK035993.1 GI:26084949
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
1
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374

		Mus musculus	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;	Carninci, P. and Hayashizaki, Y.	High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 99279253 10349636	1	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)	2	20499374 11042159	3	Shibata, K., Ito, H., Aizawa, K., Nagao, S., Sasakura, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitamura, T., Tashiro, H., Itoh, M., Sumi, R., Ishii, Y., Nakamura, S., Hazama, M., Nishigami, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuda, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multiplexed parallel sequencer Genome Res. 10 (11), 1757-1771 (2000)	4	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Ito, H., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Miyosawa, H., Kondo, S., Yananaka, I., Saio, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fieschmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikolaev, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staahl, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furumasa, A., Anoh, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hu, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marzocchi, L., Mashima, J., Mazzarelli, J., Mombauts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S. and Hayashizaki, Y. Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 685-690 (2001)	5	Nature 409 (6821), 685-690 (2001) 21085660 11217851	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)	6	(bases 1 to 2855)	Fukushima, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Adachi, J., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Koike, Y., Kondo, S., Konno, H., Kouma, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akaira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
--	--	--------------	--	----------------------------------	--	---	---	---	----------------------	---	---	---	--	---	---	--	---	-------------------	--


```

QY 1618 CGGCTTGATGCTTGTGGCCCGCCAGCGCGGAGCACCGCGGGATGGT 1664
Db 1715 CGGCTGATGCTTGTGGCCCATGCGAGCGAGCACCGTGGTGAGT 1761

RESULT 3
CB527136
LOCUS CB527136 704 bp mRNA linear EST 28-MAR-2003
DEFINITION UI-M-FY0-cfh-1-24-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
IMAGE: 6849601 5', mRNA sequence.
CB527136
VERSION CB527136.1 GI:29360609
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 704)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pyx-5.
Location/Qualifiers
1. 704
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FY0"
/note="Organ: Brain; Vector: pyx- Asc; Site_1: EcoR I;
Site_2: Not I; The library was constructed according
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pyx-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCGAGACAG. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."
BASE COUNT 149 a 180 c 212 g 163 t
ORIGIN
Query Match 27.78; Score 645; DB 14; Length 704;
Best Local Similarity 95.6%; Pred. No. 7.2e-153;
Matches 674; Conservative 0; Mismatches 30; Indels 1; Gaps 1;

QY 305 CCCGAAGGATCGACTGGTGTACTGACACTCACACAGAACTGCGAGGAAGGCA 364
Db 1 CTCGAAGGATCGACTGGTGTGCGGAGACTCACAGACAACTGTAGGAAGGCA 60

QY 365 AGAAGAGGACGAATGCACAATTTATCCAGATTCGCGCAATGTCAATGCCCTCTCACC 424
Db 61 AGAAGAGGACGAATGCACAATTTATCCAGATTCGCGCAATGCCCTCTCACC 120

```

```

QY 425 TCCTCAGCTGGGCGACCTTCGCTTTTGTATCGGAAGTCGGGGTTATTGATGTGTCCAGTT 484
Db 121 TCCTCAGCTGTGGCACCTTCGCTTTTGTATCGGAAGTCGGGGTTATTGATGTGTCCAGTT 180

QY 485 TCCAGCAGGTTGAAAGACTTGAGAGCGCGGGGAAATGTCCTTTTGGAGCCAGCTCAAC 544
Db 181 TCCAGCAGGTTGAAAGACTTGAGAGTCGCCGGGGAATGTCCCTTTTGGAGCCAGCTCAAC 240

QY 545 GGTTCAGCAGCTGTAAATGGCTGGGGCGCTCTCTACACCGCCACTGTGAAGAACTTCCTGG 604
Db 241 GGTTCAGCAGCTGTAAATGGCTGGGGCGCTCTCTACACCGCCACTGTGAAGAACTTCCTGG 300

QY 605 GGAATGAGCCCATCATCTCCGAGCTGTGGGTCCAGCTGAGGACTGGATTCGAACAGAGA 664
Db 301 GGACAGAGCCGATTTATCTCCGAGCTGTGGGTCCAGCTGAGGACTGGATTCGAACAGAGA 360

QY 665 CTTGTGCATCTGCTGCTTAATGCTCCAGCTTTGTTCGAGCTATGCTCTGAGCCAGCTG 724
Db 361 CTTGTGCATCTGCTGCTTAATGCTCCAGCTTTGTTCGAGCTATGCTCTGAGCCAGCTG 420

QY 725 AGTGGGGGATGAAGATGGAGACGATGAAATCTTTTCTTCCAGGAGACCTCCCGAG 784
Db 421 AGTGGGGGATGAAGATGGAGACGATGAAATCTTTTCTTCCAGGAGACCTCCCGAG 480

QY 785 TGTGGGACTCTATGAGCGCATCAAGTCCCAAGATGGCCGAGTGTGCGGGGAGCC 844
Db 481 TGTGGGACTCTATGAGCGCATCAAGTCCCAAGTGGCCGAGTGTGCGGGGAGCC 540

QY 845 TTGGGGCGAGGAACCTTTCAGCAGAGATGGAGCGTTCCTCAAGGCTGACCTGCTGT 904
Db 541 TTGGGGCGAGGAACCTTTCAGCAGAGATGGAGCGTTCCTCAAGGCTGACCTGCTGT 600

QY 905 GCCCAGGGCCGAGCATGGCCGCTCCGCGCTTCGAGGCTATGCGAGCTATGCGAGCTCCGG 964
Db 601 GTCCAGGGCCGAGCATGGAAAGGCTTCGCGGGTCTTCGAGGATATCACAGCTTCGAC 660

QY 965 CTCAGCCTGGAGCGGGAACCCCATCTTTATGGATCTTTTCCT 1009
Db 661 CTCAGCCTGGCGGGGAGCCCTCCTTTTATGGATCTTTTCCT 704

RESULT 4
LOCUS BM944530 691 bp mRNA linear EST 14-MAR-2002
DEFINITION UI-M-EH0p-bvr-f-24-0-UI.r1 NIH_BMAP_EH0p Mus musculus cDNA clone
IMAGE:5695895 5', mRNA sequence.
ACCESSION BM944530
VERSION BM944530.1 GI:19428115
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 691)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pyx-5.
Location/Qualifiers
1. 691
/organism="Mus musculus"
/mol_type="mRNA"

FEATURES
source

```

```
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5695895"
/tissue_type="whole brain"
/dev_stage="embryo 18.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH-BMAP_EH0p"
/note="Organ: brain; Vector: pYX-Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is CAGCCAGGAC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."
```

BASE COUNT 145 a 177 c 207 g 161 t 1 others

Query Match 27.1%; Score 632.6; DB 12; Length 691;
Best Local Similarity 95.5%; Pred. No. 1e-149;
Matches 661; Conservative 0; Mismatches 30; Indels 1; Gaps 1;

QY 318 CTGGATGCTACTGAGACTCACAGACAGAACTGCAGGAAGAAAGCAGAAAGAGGACGA 377
DB 1 CTGGATGCTGCGGAGACTCACAGACAGAACTGTAGGAAGAAAGCAGAAAGAGGACGA 60

QY 378 ATGTACAAATTTATCCAGATTCGCGCAATTTGTAATGCCCTCACCTCCACGTGGG 437
DB 61 ATGTACAAATTTATCCAGATTCGCGCAATTTGTAATGCCCTCACCTCCACGTGGG 120

QY 438 CACCTTCGCTTTGATCCGAAGTGGGGGTTATTGATGTGCCAGTTTCCAGCAGGTTGA 497
DB 121 CACCTTCGCTTTGATCCGAAGTGGGGGTTATTGATGTGCCAGTTTCCAGCAGGTTGA 180

QY 498 AAGACTTCAGAGCGCGCGGGAATGTCCTTTGAGCCAGCTCAACGGTCAGCAGCTGT 557
DB 181 AAGACTTCAGAGTGGCGGGGGAATGTCCTTTGAGCCAGCTCAACGGTCAGCAGCTGT 240

QY 558 AATGGCTGGGGCGCTCTTACACCGCCACGTGTGAAGAACTTCCTGGGACTGAGCCCAT 617
DB 241 AATGGCTGGGGCGCTCTTACACCGCCACGTGTGAAGAACTTCCTGGGACAGACCGAT 300

QY 618 CATCTCCGAGCTGTGGGTGAGCTGAGGACTGGATTCGAACAGAGACCTTTGTCATCTG 677
DB 301 TATCTCCGAGCTGTGGGTGAGCTGAGGACTGGATTCGAACAGAGACCTTTGTCATCTG 360

QY 678 GCTTAATCTCCAGCTTTGTCGAGCTATGGTCTGAGCCAGCTGAGTGGGGGATGA 737
DB 361 GCTTAATCTCCAGCTTTGTCGAGCTATGGTCTGAGCCGCGCTGAGTGGGGGATGA 420

QY 738 AGATGGAGACGATGAATCTTTTCTTTCACGAGACCTTCCGAGTGTGGACTCTGA 797
DB 421 AGATGGAGACGATGAATCTTTTCTTTCACGAGACCTTCCGAGTGTGGACTCTGA 480

QY 798 TGAGCGCATCAAGTCCCAAGATGGCCGAGTGTGTGCGGGGACCTTGGGGGACGAA 857
DB 481 TGAGCGCATCAAGTCCCAAGATGGCCGAGTGTGTGCGGGGACCTTGGGGGACGAA 540

QY 858 GACCTTCAGCAGAGATGGAGCAGCTTTCGAAGCTGACCTGCTGCTGCCAGGCCCGA 917
DB 541 GACCTTCAGCAGAGATGGAGCAGCTTTCGAAGCTGACCTGCTGCTGCCAGGCCCGA 600

QY 918 GCATGGCGCGCCCTCCGGGGTTCTCGAGGCTATGGCAGAGCTTCGGGCTCAGCCTGGAGC 977
DB 601 GCATGGAGGGCCCTCCGGGGTTCTGCAGGATATGACAGAGCTTCGACCTCAGCCTGGCGC 660

QY 978 GGGAAACCCCATCTTTTATGGGATCTTTTCCT 1009
DB 661 GGGGACCCCT-TCITTTATGGCATCTTTTCCT 691

RESULT 5
BG323723 868 bp mRNA linear EST 27-FEB-2001
LOCUS 602421820F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4560028 5',
DEFINITION mRNA sequence.
ACCESSION BG323723
VERSION BG323723.1 GI:13130160
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 868)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CMI268 row: g column: 05
High quality sequence stop: 860.

FEATURES
Location/Qualifiers
1..868
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4560028"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_14"
/note="Organ: kidney; Vector: pOT8; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using 2AP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 167 a 240 c 273 g .188 t

ORIGIN
Query Match 26.1%; Score 608; DB 10; Length 868;
Best Local Similarity 84.0%; Pred. No. 2e-143;
Matches 712; Conservative 0; Mismatches 130; Indels 6; Gaps 2;

QY 1484 TGAATTGTACACGATTGGCTCTCTGGTGGCTCCCATCTAGGTGACACAAAGTGAACA 1543
DB 2 TGAATTGTACACAGCTGGCTCTCTGGTGGCTCCGCTACTGAGTGACACAAAGTGAATA 61

QY 1544 CCAGCACTGTGGCGTCTCCAGAGCTGCTCGGAGTGTATCTTGGCCAGGACCCGCTGT 1603
DB 62 CAACCACTGTGGCGTCTCCAGAGCTGCTCAGAGTGTCTCTGGCCAGGACCCAGTCT 121

QY 1604 GCGCTTGAGCTTCGGGCTTGTGTTGGCCAGCGGAGCACCAGCGGATGG 1663
DB 122 GTGCTTGAGCTTCGGGCTGGATGAGTGTGTGGCCATGCCGGGAGCACCAGGTTGG 181

QY 1664 TTCAAGATATAGATCAGCGGATGTCTCTTTTGTGTCCAAAAGCAATCTGGAGAACATC 1723
DB 182 TCCAAGACATAGATCAGCAGATGTCTCTCTTTGTGTCTTAAAGAGCCTGAGAACATC 241

QY 1724 CCGTAGTGTGTTGAGAGTCCGGTGGCTACTGTGGGCCAGCTGGTCTGCATGTTCCCCCA 1783


```

|||||
600 GCAGCGCGCGGACAGAGGGAGCTTCTAGCTAGAGACAAGTGGGCTTAGATCTGGGGG 659
|||||
2130 TCCACCTCTCTGGGACCAAGCTATAGTCAGGAC 2163
|||||
660 CTCACCTCTCTGGGACCAAGCTATAGCCAAGAC 693

RESULT 7
LOCUS BM462308
DEFINITION BM462308.1 GI:20395053
5', mRNA sequence.
ACCESSION BM462308
VERSION BM462308.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13334 row: 1 column: 20
High quality sequence stop: 604.
Location/Qualifiers
1. .901
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6062883"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 181 a 248 c 289 g 183 t
ORIGIN
Query Match 25.1%; Score 586.2; DB 13; Length 901;
Best Local Similarity 86.3%; Pred. No. 7e-138;
Matches 648; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

1283 GCGCGCCCTCTGCTGCTACTACAGATACAGCTATCTCAGTCTGCGGCCACAGGGTGA 1342
|||||
12 GGCACCCCTCTGCTGCTACTACAGATACAGCTATCTCAGTCTGCGGCCACAGGGTGA 71
|||||
1343 CCAGCCTCTCAGGGAAGAAATATGACGCTCTACCTGGGGACAGAGGATGGACACCTCC 1402
|||||
72 CCAGCCTCTCAGGGAAGAAATATGATGTCTCTACTCTGGGGACAGAGGATGGACACCTCC 131
|||||
1403 ACCGGCTCTGCGCATTTGGAGCTCAGCTCAGTCTTCTTGAGGATCTGGCTTTGCCAG 1462
|||||
132 ACCGAGCAGTGGGATCGGAGCTCAGCTCAGCTGCTTTGAAGATCTGGCTTTATTCGCCAG 191
|||||
1463 AACCAAGCGGCTTGAGAGCATGAATTTACACGATTTGGCTCTCTGTTGGCTCCCAT 1522
|||||
192 ACCCAGACGCTTGAGACATGAATTTGACACAGCTGGCTCTCTGTTGGCTCCCGTA 251
|||||
1523 CTGAGGTGACAACTGACACCAAGCTGTGGCGCTCTCCAGAGCTGCTCGGAGTGA 1582
|||||

```

```

Db 252 CTGAGGTGACACAGTGAATACAACTGTGTGGCGCTCTCCAGAGCTGTCTCAGAGTGA 311
Qy 1583 TCCTGGCCAGGACCCCGTGTGGCCCTGGAGCTTCCGGCTTGTGCTGGCCACG 1642
|||||
Db 312 TCCTGGCCAGGACCCAGTCTGTGGCTGGAGCTTCCGGCTGGATGCTGTGGCCCATG 371
|||||
Qy 1643 CCGCGAGCAGCCGGGATGTTCAAGATATAGAGTACAGCGGATGCTCTCTTTTGTGTC 1702
|||||
Db 372 CCGGGAGCAGCCGAGGTTGTTCCAGACATAGAGTCAGAGATGTCTCTCTTTGTGTC 431
|||||
Qy 1703 CAAAAGAACCTGGAGACATCCCGTACTGTTTGAAGTTCGGTGGCTACTGTGGCCACG 1762
|||||
Db 432 CTAAGAGCCTGGAGAACGTCACAGTGTGTTGAAGTTCGGTGGCTACTGTGGCGATG 491
|||||
Qy 1763 TGGTCTGCTGCTATGCTTCCCGAGTTCGCTGGGATCTGTGTGGCAGCAGCCAGTG 1822
|||||
Db 492 TGGTCTGCTGCTATGCTTCCCAAGCTCAGCATGGGATCTGTGTGGCAGCAGCCAGTG 551
|||||
Qy 1823 GAGTACTGCTGCTACTTCCCGGAGGATGAGTATAGAGTGGTGGTGACCCAGGGCCA 1882
|||||
Db 552 GAGTACTGCTGCTACTTCCCGGAGGATGAGTATAGAGTGGTGGTGACCCAGGGCCA 611
|||||
Qy 1883 TGGGGCTTATGCTTGGAGTGTGAGAGGTTGAGCGCCCGCTGTGTGGCAGCAGCTTATA 1942
|||||
Db 612 TGGGGCTTATGCTTGTGATGTCAAGAGGTTGGGCGAGCCATGTGTGTAGCAGCTTACA 671
|||||
Qy 1943 GCTTGTGTGGGCGAGCAGCGGGGACCTCAAAACCGGGCCACACCGTTGTGGGGCTG 2002
|||||
Db 672 GCTTGTGTGGGCGAGCAGCAGATGCTCCGAGCGGGCCACACACATGTGGGGCGGG 731
|||||
Qy 2003 GATTGTTGGCTTCTCTCTGGGTTCTTTCG 2033
|||||
Db 732 ACTGGCTGGGCTTCTTCTTCTGGGGATCTCTCC 762

RESULT 8
LOCUS BM462308
DEFINITION BM462308.1 GI:18511348
5', mRNA sequence.
ACCESSION BM462308
VERSION BM462308.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12111 row: p column: 14
High quality sequence stop: 691.
Location/Qualifiers
1. .1154
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5491453"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_67"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr.

```

Average insert size 1.75 kb. Library constructed by Life Technologies."	
BASE COUNT	221 a 380 c 315 g 238 t

[illegible]

RESULT	9
AW953466	
LOCUS	
DEFINITION	733 bp mRNA linear EST365536 MAGE resequences, MAGB Homo sapiens cDNA, mRNA sequence.
ACCESSION	AW953466
VERSION	GI:8143149
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	

REFERENCE	1 (bases 1 to 733)
AUTHORS	Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and Quackenbush,J.
TITLE	Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray
JOURNAL	Unpublished
COMMENT	Contact: John Quackenbush The Institute for Genomic 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 3528 Fax: 301 838 0208 Email: johnq@tigr.org Plate: 51
FEATURES	Seq primer: Reverse.
SOURCE	Location/Qualifiers 1..733 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone_lib="MAGE sequences, MAGB" /note="Vector: pBluescriptSkm"
BASE COUNT	133 a 207 c 238 g 154 t
ORIGIN	1 others

Query Match	23.9%	Score 557.6	DB 10	Length 733
Best Local Similarity	86.5%	Pred. No. 1.2e-130		
Matches 627	Conservative 0	Mismatches 95	Indels 3	Gaps 1
QY	1485	GAATTTGTACCAAGATTGGCTCTCGTGGGCTCCCATACTGAGTGGACACAAGTGAACAC	1544	
Db	11	GAATTTGTACCAAGCTGGCTCTCGTGGTCTCCCGTACTGAGTGGACACAAGTGAATAC	70	
QY	1545	CAGCAACTGTGGCGCTCTCCAGAGCTGCTCGGAGTGTATCTGCGCCAGGACCCCGTGTG	1604	
Db	71	AACCAACTGTGGCGCTTCAGAGCTGCTCAGAGTGCATCCTGGCCGAGACCAAGTCTG	130	
QY	1605	CGCTGGAGTTCGGGCTTGATCTTGTGTGGCCACGCGCGGAGACACCGCGGATGGT	1664	
Db	131	TGCTTGAGCTTCGGCTTGATGAGTGTGTGGCCATGCGGGGAGCACCGAGGGTTGGT	190	
QY	1665	TCAAGATATAGATCAGCGGATCTCTTCTTTGTCTCAAAGAAGCACTGGAGACATCC	1724	
Db	191	CCAAGACATAGATCAGCAGATCTCTCTTTGTCTTAAAGAGCTCGAGAACGTCC	250	
QY	1725	CGTAGTGTTTGAAGTTCGGTGGCTACTGTGGCCAGCTGCTCTGCATGTTTCCCCCGAG	1784	
Db	251	AGTAGTGTTTGAAGTTCGGTGGCTACAGCTGGCATGTGGTCTTGCCATGTTCTCCAG	310	
QY	1785	TTCTGCGTGGGCATCTGTGTGGCACCAAGCCAGTGGAGTGAAGTCTCACTCCCGG	1844	
Db	311	CTCAGCATGGCATCCTGTGTGTGGCACCAAGCCAGTGGAGTGAAGTCTCACTCA	370	
QY	1845	GAGGGATGGAAGTGTGTGTGACCCACAGGGGCCATGGGGCTTATGCTTGCAGAGT	1904	
Db	371	GCGGGATGGACTGGAGTGTGTGTGACCCACAGGGGCCATGGGGGCTTATGCCCTGTAATG	430	
QY	1905	TCAGGAGGTGGAGCGCCCGCTGGTGGCTGCTTATAGCTTGGTGTGGGGCAGCCAGCG	1964	
Db	431	TCAGGAGGTGGGGCAGCCCATGTGTGTAGCAGCTTACAGCTTGGTATGGGGCAGCCAGCG	490	
QY	1965	GGCAGCCTCAAAACGGGGCCACACCGTTGTGGGGGTGGATGTTGGCTTCTTCCTGGG	2024	
Db	491	AGATGCTCCGAGCGGGGCCACAC---	547	
QY	2025	TGTTCTTGACGATCCCTCACTCTCTCTCTGATTGTGCGCCGTACAGACGCTCGGGACA	2084	
Db	548	GATTCTCGAGCATCCCTGACTCTCTCTCTGATTGTGCGCGTACAGACGACGGCGACA	607	
QY	2085	GAGGGAGCTTCTAGCTAGACACAAGTGGGCTTAGATCTGGGGGCTCCACCTTCTGGGAC	2144	
Db	608	GAGGGAACCTTCTGGCTANAGACAAGTGGCCCTGGACCTGGGGCTCCACCTTCTGGGAC	667	

[illegible]

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 701)
NH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished
Contact: Robert Strausberg, Ph.D. Email: c9apbs-re@mail.nih.gov Tissue Procurement: Dr. Jim Lin, University of Iowa CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov This clone was contributed by the Brain Molecular Anatomy Project (BMAP) The following repetitive elements were found in this CDNA sequence: 116-149, >G-rich#low_complexity 166-207, >(CAG)nSimple_repeat (matched compliment) Seq primer: pVX-5.
Location/Qualifiers
1..701
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6834138"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DHI0B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FY0"
/note="Organ: Brain; Vector: pVX- Asc; Site_1: Ecot R I; Site_2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cdna synthesis was primed with oligo-df primer containing a Not i site. Double strand cdna was size selected according to mRNA size fraction, ligated with Ecot R adaptor, digested with NotI and then cloned directionally into pVX-Asc vector. The library tag sequece located between the Not i site and the polyA tail is ACGGAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP); "Gene Discovery in the Developing Mouse Nervous System", supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."
BASE COUNT 123 a 230 c 205 g 143 t
ORIGIN
Query Match 23.5%; Score 540; DB 14; Length 701;
Best Local Similarity 95.3%; Pred. No. 3.6e-126;
Matches 568; Conservative 0; Mismatches 25; Indels 3; Gaps 1;
QY 1 ATGCTTTTCAGCGCCGAGCGGCCCGCCCCGCGGCGGTCTATTAACGCCCGCGGTCTGCGGC 117
Db 106 ATGCTTGCGCAGGGCGAGCGGCCGCCGCGGCGCGGCGGTCTATTAACGCCCGCGGTCTGCTTC 165
QY 61 CGCGCGCGCGCTGTC---GCTGCTGCTGCTGCGCATACTAAACGCCCGCGGTCTGCGGC 117
Db 166 CGCGCGCGCT 235
QY 118 CGCGTCCCCTCAAGTGCACCAGAACCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 177
Db 226 CCGCTCCCCCGCTCGGTGCCAGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 285
QY 178 ACCCGGTTTTGACGGCTCTCATACGTACAATTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 237
Db 286 ACCCGGTTTTGCGGCGCCCTCATACGTACAATTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 345
QY 238 CACACACTTTAGTTCGGTGCACGGGATAGCATCTTCGCTTTAAACCTTCCCTTCTCTCTCTGCTGCTGCT 297
Db 346 CACACACTTTAGTTCGGTGCACGGGATAGCATCTTCGCTTTAAACCTTCCCTTCTCTCTCTGCTGCTGCT 405

```
QY 298 GAAAGACCCGGAAGATCAGCTGGATGGTACCTGAGACTCAGACAGAGAACTGCAGGAAG 357
|||||
Db 406 GAAAGCCCTCGAAGATCAGCTGGATGGTGGCCGAGACTCAGACAGAGAACTGTAGGAAG 465

QY 358 AAAGGCAACAAGAGAGAGCAATGTCCAAATTTTATCCAGATTCTCGCCATTGTCAATGCC 417
|||||
Db 466 AAAGGCAACAAGAGAGAGCAATGTCCAAATTTTATCCAGATTCTCGCCATTGTCAATGCC 525

QY 418 TCTACCTCTCTCACGTGCGGCACCTTCCGCTTTTGTATCCGAAGTGGGGGTTATTGATGTG 477
|||||
Db 526 TCTACCTCTCTCACGTGCGGCACCTTCCGCTTTTGTATCCGAAGTGGGGGTTATTGATGTG 585

QY 478 TCCAGTTTCCAGCAGCTTGAAGAGACTTGAGAGCGCGCGGGGAAATGTCTTTTGAAGCCA 537
|||||
Db 586 TCCAGTTTCCAGCAGCTTGAAGAGACTTGAGAGCGCGCGGGGAAATGTCTTTTGAAGCCA 645

QY 538 GCTCAACGGTCAGCAGCTGAATGGTGGGGCGTCTCTACACGCCCACTGTGAA 593
|||||
Db 646 GCTCAACGGTCAGCAGCTGAATGGTGGGGCGTCTCTACACGCCCACTGTGAA 701

RESULT 13
BI915256
LOCUS
DEFINITION BI915256 772 bp mRNA linear EST 16-OCT-2001
603184795F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5248726 5',
mRNA sequence.
ACCESSION BI915256
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 772)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLML1627 row: n column: 23
High quality sequence stop: 770.
Location/Qualifiers
1..772
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5248726"
/lab_host="DH10B"
/clone_lib="NIH_MGC_121"
/note="Organ: brain; Vector: pCMV-Sport6; Site:1: NotI;
Site:2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH_MGC Library."
BASE COUNT 158 a 211 c 242 g 161 t
ORIGIN
```

FEATURES

```
source
CB244307
DEFINITION UI-M-FY0-cdp-d-18-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
IMAGE: 6832579 5', mRNA sequence.
ACCESSION CB244307
VERSION
KEYWORDS EST.
SOURCE CB244307.1 GI:28365951
ORGANISM Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 758)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
```


Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
e mouse tissues.

FEATURES

source

Location/Qualifiers

1..669
/organism="Mus musculus"
/mol_type="mrna"
/db_xref="taxon:10090"
/clone="9630025H20"
/tissue_type="cerebellum"
/dev_stage="16 days neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 16 days neonate cerebellum"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGAGATCCCAAGAGCTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 370.0. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGAGATCTCGAGTAAATTAATCCGCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pbluescript KS(+) after bulk excision from Lambda FIC I."

BASE COUNT 119 a 220 c 194 g 136 t
ORIGIN

Query Match 20.8%; Score 484; DB 10; Length 669;

Best Local Similarity 94.7%; Pred. No. 6.1e-112;

Matches 534; Conservative 0; Mismatches 25; Indels 5; Gaps 3;

QY	1	ATGCTTGGCAGGGCGAGCGGCCCGCGGGCCCGCGCGCCCTCCGGTCTTTCCCTTC	60
Db	108	ATGCTGGCGAGGGCGAGCGGCCCGCGGGCCCGCGCGCCCTCCGGTCTTTCCCTTC	167
QY	61	CGCGCGCGCGTGC---GCTGCTGCTGCTGGCGATACAAAGCGCCCGGGTGGCGC	117
Db	168	CGCGCGCGCGTGCCTGCTGCTGCTGGCGATGCTGAGCGCCCGGGTGGCGC	227
QY	118	CGCGTCCCGCGCTCAGTCCAGAACCTCGCTGCGCCATCTCGAGGCTGACTCCTATCTC	177
Db	228	CGCGTCCCGCGCTCGGTGCCAGAACCTCGCTGCGCCATCTCGAGGCTGACTCCTATCTC	287
QY	178	ACCGGTTTGCAGCGTCTCATACGACAAATTAATCTGCTCTCTCTGATCCTGCCCTCC	237
Db	288	ACCGGTTTGGCGCCCTCATACGACAAATTAATCTGCTCTCTCTGATCCTGCCCTCC	347
QY	238	CACACACTTACGTCGGTGCACGGGATAGCATCTTCGCTTTAAACCTCCCTCTCTGGG	297
Db	348	CACACACTTACGTCGGTGCACGGGATAGCATCTTCGCTTTAAACCTCCCTCTCTGGG	407
QY	298	GAAGACCCCGAAGGATCGACTGGATGGTACTGAGACTCACAGACAGAACTGCAGGAAG	357
Db	408	GAAGACCTCGAAGGATCGACTGGATGGTCCCGAGACTCACAGACAGAACTGTAGGAAG	467
QY	358	AAAGCAAGAAAGAGGACAAATGTCAAAATTTATCCAGATTCCTGCCATTGTCAATGCC	417
Db	468	AAAGCAAGAAAGAGGACAAATGTCAAAATTTATCCAGATTCCTGCCATTGTCAATGCC	527
QY	418	TCTACCTTCCACGTCGGGACCTTCGCTTTTGGATCCGAAGTGGGGTTATTGATGTG	477
Db	528	TCTACCTTCCACGTCGGGACCTTCGCTTTTGGATCCGAAGTGGGGTTATTGATGTG	587
QY	478	TCCAGTTTCCACGTTTGAAGACTTGAGAGCGCGCGGGGAAATGTCTTTTGGAGCA	537
Db	588	TCCAGTTTCCACGTTTGAAGACTTGAGAGCTTGAGAGTGCC--GGGAAATGTCTTTGAGCA	646

QY 538 GCTCAACGGTCTCAGCAGCTGTAATG 561
||||| |||||||||
Db 647 GCTCAAC-GTCAGCAGCTGTAATG 669

Search completed: September 26, 2003, 05:40:49
Job time : 4617.83 secs